ChemoCentryx, Inc
Protocol #: CL016_168

A Randomized, Double-Blind, placebo-Controlled, Parallel Group, Phase 2 Study to Evaluate the Safety and Efficacy of Avacopan in Subjects with Moderate to Severe Hidradenitis Suppurativa

Statistical Analysis Plan

Version 7.0
Document last saved: 1-Dec-2020 10:10 AM

Prepared for:
ChemoCentryx, Inc.
Mountain View, California 94043, USA
Phone: +1 (650) 210-2900

Prepared by:
PPD
PPD

LIST OF ABBREVIATIONS AND ACRONYMS

AAV	ANCA-associated vasculitis
AE	adverse event
AESI	adverse event of special interest
ALT	alanine aminotransferase
AN	abscess and inflammatory nodule (count)
ANA	antinuclear antibodies
ANC	absolute neutrophil count
ANCA	anti-neutrophil cytoplasmic autoantibody
ANCOVA	analysis of covariance
AST	aspartate aminotransferase
AT	Aminotransferase
ATC	Anatomic Therapeutic Chemistry
AUC	area under the plasma concentration-time curve
b.i.d.	twice daily
C5a	complement component 5a fragment
C5aR	receptor for C5a
C5b-9	membrane attack complex or terminal complement complex
CA	competent authority
C_{max}	maximum (maximal) plasma concentration
C_{min}	minimum plasma concentration
CMH	Cochran-Mantel-Haenszel
COVID-19	Coronavirus Disease 2019
CPK	creatine phosphokinase
CRA	Clinical Research Associate
CRF	Case Report Form
CRO	clinical research organization
CSR	clinical study report
CYP3A4	cytochrome P450 3A4
DLQI	Dermatology Life Quality Index
DMC	data monitoring committee
ECG	Electrocardiogram
eCRF	electronic case report form
EDC	electronic data capture
EQ-5D-5L	EuroQuality of Life-5 Domains-5 Levels
FDA	Food and Drug Administration
g	Gram
HBV	hepatitis B virus
HCV	hepatitis C virus
HiSCR	hidradenitis suppurativa clinical response
HiSQOL	Hidradenitis Suppurativa Quality of Life
HIV	human immunodeficiency virus
HS	hidradenitis suppurativa
HS-PGA	Hidradenitis Suppurativa-Physician Global Assessment
Ig	Immunoglobulin
IgAN	immunoglobulin A nephropathy

IGRA	interferon γ release assay
IHS4	International HS Severity Scoring System
IND	Investigational New Drug
INR	International normalized ratio
IRT	interactive response technology
ITT	intent-to-treat
IV	intravenous
kDa	Kilodalton
kg	Kilogram
LOR	loss of response
MedDRA	Medical Dictionary for Regulatory Activities
Mg	Milligram
mL	Milliliter
MPO	myeloperoxidase
MSS	modified Sartorius score
N	Number
NF	National Formulary
NRI	Non-Responder Imputation
NRS	numeric rating scale
NRS30	proportion of subjects achieving at least 30% reduction and at least 1 unit reduction from baseline in the subject's global assessment of skin pain on a numeric rating scale
PK	pharmacokinetic(s)
PP	Per-protocol
PT	prothrombin time
QA	quality assurance
QC	quality control
QT/QTc	Q-T interval on ECG; corrected Q-T interval
RBC	red blood cell(s)
RNA	ribonucleic acid
SAE	serious adverse event
SAP	statistical analysis plan
SAR	serious adverse reaction
SD	standard deviation
SEM	standard error of the mean
SF-36 v2	Short Form-36 version 2.0
SOP	standard operating procedure
SUSAR	suspected unexpected serious adverse reaction
TB	tuberculosis
TBili	total bilirubin
TEAE	Treatment emergent adverse event
ULN	upper limit of normal
WBC	white blood cell
WHODD	World Health Organization Drug Dictionary
WPAI:SHP	Work Productivity and Activity Impairment Questionnaire: Specific Health Problem

Contents

I. Introduction

A. Background

The purpose of this document is to provide a description of the statistical methods and procedures to be implemented for the analysis of data from ChemoCentryx, Inc. Protocol CL016_168. This document is based on Protocol Amendment 2.0 (31JULY2019). If circumstances arise during the study such that more appropriate analytic procedures become available, the statistical analysis plan (SAP) may be revised. The statistical definitions and analytical methods described in this SAP supersede that in the protocol. Any revisions to the primary endpoint analyses and significant revisions to the secondary endpoint analyses will be made prior to the database freeze for the data in the double-blind treatment period through Week 12. Reasons for such revisions will be described in the final Clinical Study Report (CSR).

B. Protocol and Amendment History

This Statistical Analysis Plan (SAP) is based on 31JUL2019 of Protocol CL016_168, Amendment 2.0.

Version	Approval Date
Original Protocol	19 October 2018
Amendment 1	19 March 2019
Amendment 2	31 July 2019

II. Protocol Objectives

The study is a randomized, double-blind, placebo-controlled, three-group Phase 2 trial in approximately 390 subjects with moderate to severe hidradenitis suppurativa (Hurley Stage II or III). Subjects will be randomized 1:1:1 to a treatment of 10 mg avacopan twice daily, 30 mg avacopan twice daily or placebo for 12 weeks. Subjects treated with 10 mg or 30 mg twice daily during the blinded, placebo-controlled 12-week treatment period (Period 1, day 1 to week 12) will be followed by an additional 24-week, active treatment period (Period 2, week 12 to week 44) during which they will continue to receive the same dose regimen, either 10 mg or 30 mg avacopan twice daily.

A. Primary

1) Evaluation of the efficacy of avacopan compared to placebo in subjects with Hurley Stage II or III hidradenitis suppurativa (HS) based on subjects achieving a Hidradenitis Suppurativa Clinical Response (HiSCR) after 12 weeks of treatment. HiSCR is defined as at least a 50% reduction in abscess and inflammatory nodule (AN) count and no increase in abscess count and no increase in draining fistula count at Week 12 relative to baseline.

2) Evaluation of the safety of avacopan compared to placebo in these subjects based on the adverse event incidence, changes from baseline in laboratory parameters, and vital signs

B. Secondary

1) Evaluation of the efficacy of avacopan compared to placebo in these subjects include:
    a. The subject's global assessment of skin pain numeric rating scale (NRS),
    b. The modified Sartorius score, and
    c. Achieving an AN count of 0, 1, or 2.

2) Assessment of subject-reported outcomes including health-related quality-of-life changes based on the Short Form-36 version 2 (SF-36 v2), the EuroQOL-5D-5L (EQ-5D-5L), the Hidradenitis Suppurativa Quality of Life (HiSQOL) Index, the Dermatology Life Quality Index (DLQI), and the Work Productivity and Activity Impairment Questionnaire: Specific Health Problem (WPAI:SHP) with avacopan compared to placebo.
3) Evaluation of the pharmacokinetic profile of avacopan in subjects with HS.
4) Evaluation of the safety and efficacy of avacopan treatment from Day 1 by each timepoint up to Week 44 in subjects with HS.
5) Evaluation of the efficacy of avacopan compared to placebo in these subjects include:
    a. The Sartorius score, International HS Severity Scoring System (IHS4) score, HS Physician Global Assessment (HS-PGA),
    b. Proportion of subjects who experienced flare, who experienced loss of response during Period 2, who received oral antibiotic rescue therapy or lesion intervention, and who received disallowed opioid pain therapy, and
    c. The duration of flare in days.
6) Evaluation of health-economic information.

III. Study Endpoints

A. Efficacy Endpoints

- **Primary Efficacy Endpoint**

The primary efficacy endpoint is the proportion of subjects achieving Hidradenitis Suppurativa Clinical Response (HiSCR) at Week 12. A response is defined as a reduction of at least 50% in abscess and inflammatory nodule count with no increase in abscess count and no increase in draining fistula count compared to baseline.

- **Secondary Efficacy Endpoints**

The following secondary efficacy endpoints will be analyzed according to the rank order as the follows:

1. Change from Baseline in total AN count at Week 12;
2. Proportion of subjects achieving at least 30% reduction and at least 1 unit reduction from Baseline in the subject's global assessment of skin pain (NRS30) in subjects with a Baseline NRS of at least 3, evaluated at Week 12; weekly averages of daily pain will be calculated based on subjects' daily diary recording of the worst pain experienced in the previous 24 hours.
3. Proportion of subjects with baseline Hurley Stage II who achieved an abscess and inflammatory nodule count of 0, 1, or 2 at Week 12.
4. Reduction of IHS4 score relative to baseline at Week 12.
5. Change from Baseline in inflammatory nodule count at Week 12.
6. Change from Baseline in abscess count at Week 12.
7. Change from Baseline in draining fistula count at Week 12.
8. Change from Baseline to Week 12 in the modified Sartorius score to quantify the severity change of HS;

- **Other Efficacy Endpoints**

The following efficacy endpoints will be analyzed from Baseline to each timepoint up to Week 44, where applicable:

1. Proportion of subjects achieving HiSCR, $HiSCR_{75}$, and $HiSCR_{90}$;
2. Proportion of subjects achieving at least 30% reduction and at least 1 unit reduction from Baseline in the subject's global assessment of skin pain (NRS30), in subjects with a Baseline NRS of at least 3;
3. Proportion of subjects with baseline Hurley Stage II who achieved an abscess and inflammatory nodule count of 0, 1, or 2;
4. Change from Baseline in inflammatory nodule count, abscess count, non-inflammatory count, draining fistula, non-draining fistula, total fistula count, and total AN count;
5. % Change from Baseline in inflammatory nodule count, abscess count, non-inflammatory count, draining fistula, non-draining fistula, total fistula count, and total AN count among subjects who have at least one corresponding lesion at Baseline;
6. Change from Baseline in the Sartorius score, modified Sartorius score, IHS4 score, and HS-PGA to quantify the severity change of HS;
7. Change from Baseline in subject-reported outcomes: SF-36 v2, EQ-5D-5L, HiSQOL and DLQI;

8. Proportion of subjects who experienced flare, defined by at least a 25% increase in AN counts with a minimum increase of 2 AN lesions relative to Baseline;
9. Duration of flare in days (calculated from the day when flare is observed to the day prior to the observation that flare is no longer present; of note, there could be multiple periods that flares are observed, in which case, the total days from the multiple periods will be used);
10. Proportion of subjects who experience at least 25% increase in draining fistula counts with a minimum increase of 2 draining fistula counts relative to Baseline;
11. During Period 2, proportion of subjects with a loss of response, (LOR) defined as loss of at least 50% of AN count improvement achieved from Week 12 in Period 1;
12. Time to LOR during Period 2;
13. Proportion of subjects who received add-on antibiotic or other therapy during Period 1 and Period 2 including the follow-up period separately;
14. Proportion of subjects who start disallowed opioid pain therapy during Period 1 and Period 2 including the follow-up period separately;
15. Proportion of subjects who undergo lesion intervention due to HS during Period 1 and Period 2 including the follow-up period separately;
16. Number of lesion interventions due to HS during Period 1 and Period 2 including the follow-up period separately;
17. Health-economic information:
    a. WPAI:SHP: Change from Baseline to each timepoint during Period 1 and Change from Week 12 to each timepoint during Period 2 and including the follow-up period;
    b. Hospitalizations (cumulative): Number of Hospitalizations total and due to HS, Number of days hospitalized, Number of days missed from work during Period 1 and Period 2 including the follow-up period separately;
    c. Emergency or Urgent Care visits (cumulative): Number of visits total and due to HS, Number of days (or hours where applicable) missed from work during Period 1 and Period 2 including the follow-up period separately;
    d. Lesion interventions for HS: Number of days (or hours where applicable) missed from work due to HS lesion interventions during Period 1 and Period 2 including the follow-up period separately.

B. Safety Endpoints

1. Subject incidence of treatment-emergent serious adverse events, adverse events, and withdrawals due to adverse events;
2. Change from Baseline and shifts from Baseline in all safety laboratory parameters;
3. Change from Baseline in vital signs and significant changes in physical examination abnormalities.

IV. Study Design

A. Design Overview

The study is a randomized, double-blind, placebo-controlled, three-group Phase 2 trial in approximately 390 subjects with moderate to severe hidradenitis suppurativa (Hurley Stage II or III). Subjects will be randomized 1:1:1 to a treatment of 10 mg avacopan twice daily, 30 mg avacopan twice daily or placebo for 12 weeks. Other systemic treatments for HS including anti-TNF-α treatments are prohibited. Stable antibiotic therapy with doxycycline or minocycline is allowed as specified in the protocol. Subjects treated with 10 mg or 30 mg twice daily during the blinded, placebo-controlled 12-week treatment period (Period 1) will be followed by an additional 24-week, active treatment period (Period 2) during which they will continue to receive the same dose regimen, either 10 mg or 30 mg avacopan twice daily. Subjects on placebo who complete Period 1 will be re-randomized 1:1 to receive 10 mg or 30 mg avacopan twice daily in Period 2. During Period 2 the treatment assignment to 10 mg or 30 mg twice daily will not be disclosed to the subject, study site personnel or the Sponsor. Thereafter, all subjects will be followed for 8 weeks without study drug before they exit the study.

To obtain balance across treatment groups, a stratified randomization scheme will be implemented. The stratification factors and strata within each factor are listed below.

1. Hurley Stage (Stage II vs. III):
    a. Stage II disease: one or more widely separated recurrent abscesses with tract formation and scars, or
    b. Stage III disease: multiple interconnected tracts and abscesses across an entire area, with diffuse or near diffuse involvement.

2. Concomitant antibiotic therapy (Yes vs. No)
    a. Concomitantly treated with doxycycline or minocycline as the only allowed antibiotic treatment for HS, or
    b. No concomitant antibiotic therapy.

3. Anti-TNF-α treatment (Treatment naïve vs. Previous treatment):
    a. Did not previously receive anti-TNF-α drug such as adalimumab or infliximab (anti-TNF-α drug naïve), or
    b. Previously (but no longer) received an anti-TNF-α drug and
        - completed anti-TNF-α treatment but may have relapsed, or
        - was intolerant to anti-TNF-α treatment, or
        - previously failed to respond or inadequately responded to anti-TNF-α treatment.

Not more than 20% of subjects will be in stratum 2a.

For each subject, the study will consist of 4 periods:

- Screening Period (≤ 4 weeks)
- Double-Blind Treatment Period (12 weeks)
- Active Treatment Period (24 weeks)
- Follow-up Period (8 weeks)

Subjects will be screened for eligibility based on the stage of the disease and their health status. The screening period will be up to 28 days. The primary efficacy analysis will occur when the last enrolled subject has completed the Week 12 visit. After the blinded 12-week treatment period, all subjects will continue an additional 24-week treatment with either 10 mg or 30 mg avacopan twice daily. The treatment assignment to 10 mg or 30 mg twice daily will not be disclosed to the subject, study personnel at the site or the sponsor. Thereafter, all subjects will be followed for 8 weeks.

B. Study Population

Eligible adult subjects (at least 18 years of age) with moderate to severe hidradenitis suppurativa (Hurley stage II or III) on standard-of-care therapy as specified by the inclusion and exclusion criteria in the protocol Section 4, are allowed to enter the study.

C. Sample Size Estimation

The study will enroll approximately 390 subjects. In two Phase 3 clinical trials with adalimumab (PIONEER I and II; Kimball et al, 2016a), the proportions of subjects in the placebo control group achieving a HiSCR at Week 12 were 26.0% and 27.6%, respectively. Hence, if the control group only receives placebo, one may assume an HiSCR at Week 12 of approximately 30%. The average HiSCR at Week 12 in the adalimumab groups in PIONEER I and II were 41.8% and 58.9%, respectively with an average of approximately 50% (Kimball et al, 2016a). The attrition rate throughout the trial is estimated to be approximately 7%. A sample size of approximately 130 subjects per treatment group (390 in total) at a type I error rate of $\alpha = 0.05$ (two-sided) provides approximately 90% power to detect a 20% superiority of avacopan compared to the placebo control group in HiSCR at Week 12, assuming an HiSCR at Week 12 of 50% in the avacopan group.

D. Treatment Randomization

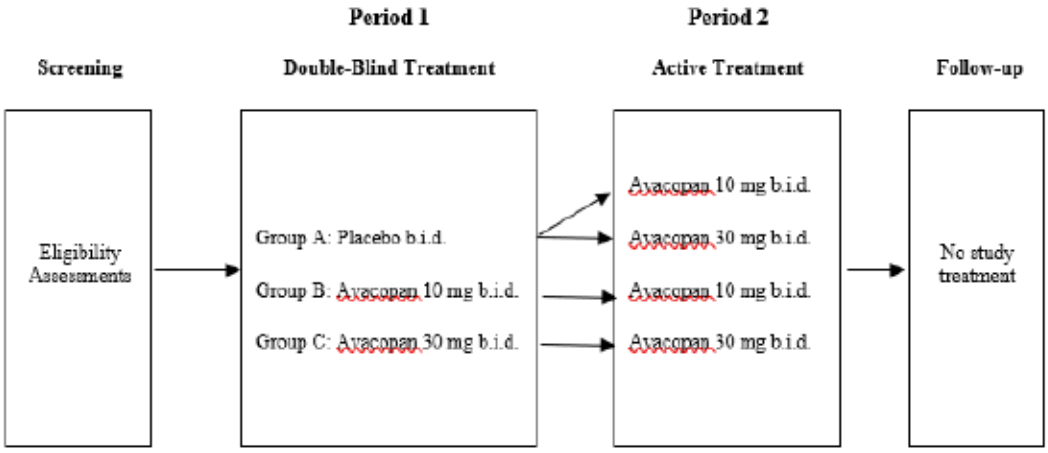
Subjects will be randomized 1:1:1 to one of the three treatment groups within each stratum based on a non-dynamic, list-based blocked randomization scheme

receive 10 mg avacopan twice daily, 30 mg avacopan twice daily or matching placebo for 12 weeks in a double-blind, placebo-controlled manner stratified by:

1) Hurley stage (Stage II vs III)
2) Concomitant antibiotic therapy (Yes vs No)
3) Anti-TNF-α treatment (Treatment naive vs Previous treatment)

E. Assessment Schedule

Subjects will undergo a Screening Period, Double-Blind Treatment Period, Active Treatment Period and Follow-up Period. Please see below for a detailed Study schedule, including all measurements and evaluations for the entire study period presented in tabular form.



Duration	≤ 28 days	84 days (12 weeks)	168 days (24 weeks)	56 days (8 weeks)
Study Visits	One or more visits to complete screening procedures	Day 1 and, Weeks 2, 4, 8 and 12	Weeks 16, 20, 24, 28, 32, 36	Week 44

Study Schedule

	Screening		Period 1 Double-Blind Treatment					Period 2 Active Treatment						Follow-up
Study Week[1]	**≤4 weeks[1]**			**2**	**4**	**8**	**12**	**16**	**20**	**24**	**28**	**32**	**36**	**44**
Study Day[1]	**≤28 days[1]**		**1**	**15**	**29**	**57**	**85**	**113**	**141**	**169**	**197**	**225**	**253**	**309**
		-7 days												
Informed consent	X													
Review inclusion/exclusion criteria	X													
Demographics, medical history, prior treatments	X													
Physical examination[2]	X		X[3]	X	X	X	X	X	X	X	X	X	X	X
Vital signs	X		X[3]	X	X	X	X	X	X	X	X	X	X	X
Serum pregnancy test for women of childbearing potential	X		X		X	X	X	X	X	X	X	X	X	
Urine pregnancy test for women of childbearing potential			X[3]											
HIV, HBV, HCV testing	X													
Screening for tuberculosis[4]	X													
12-Lead ECG	X						X							X
Serum chemistry incl. Liver Function Tests, Hematology	X		X[3]	X	X	X	X	X	X	X	X	X	X	X
Urinalysis	X		X[3]		X	X	X	X	X	X	X	X	X	X
Biomarker plasma sample collection[15]			X[15]				X						X	
Stratification and randomization			X				X[10]							
Record location and number of HS inflammatory nodules, abscesses, fistulae, and scars	X		X[3]	X	X	X	X	X	X	X	X	X	X	X
Record Hurley Stage	X													
Subject global assessment of skin pain daily diary recording[5]		X[3]→	→	→	→	→	→	→	→	→	→	→	→	→X

	Screening		Period 1 Double-Blind Treatment					Period 2 Active Treatment						Follow-up
Study Week[1]	≤4 weeks[1]			2	4	8	12	16	20	24	28	32	36	44
Study Day[1]	≤28 days[1]		1	15	29	57	85	113	141	169	197	225	253	309
		-7 days												
Record items for the Sartorius and modified Sartorius scores calculation[6]			X[3]	X	X	X	X	X	X	X	X	X	X	X
Photograph selected lesions[12]			X[12]	X	X[12]	X	X[12]	X[12]	X	X	X[12]	X	X[12]	X
Record IHS4 score[7]			X[3]	X	X	X	X	X	X	X	X	X	X	X
HS-PGA[14]			X[3]	X	X	X	X	X	X	X	X	X	X	X
SF-36 v2 and EQ-5D-5L Patient-Reported Outcomes[8]			X[3]		X		X	X			X		X	X
HiSQOL Index[8]			X[3]		X		X	X			X		X	X
Health-Economic Information[13]			X[3]	X	X	X	X	X	X	X	X	X	X	X
Study drug dispensing			X[3]				X			X				
Study drug compliance			X	X	X	X	X	X	X	X	X	X	X	
Study drug accountability[11]							X[11]			X[11]			X[11]	
PK plasma sample collection[9]			X	X	X	X	X	X	X		X		X	X
Concomitant medications	X		X	X	X	X	X	X	X	X	X	X	X	X
Adverse event assessment	X		X	X	X	X	X	X	X	X	X	X	X	X

[1]Screening period may be exceeded by up to 3 days. Week 1 through 36 visits may occur within a ± 2-day window. The Week 44 visit may occur within a ± 4-day window.

[2]Physical examination will include body weight measurements. Height will be measured only at Screening.

[3]These procedures must be performed before taking the first dose of blinded study drug (Avacopan or placebo).

[4]Screening for tuberculosis: see exclusion criteria for specifics.

[5]Subjects will record the maximum severity pain on a numeric rating scale from 0 (no skin pain) to 10 (skin pain as bad as can be imagined) in a daily diary from one week prior to Day 1 through the Week 44 visit.

[6]Twelve body areas will be evaluated to calculate the Sartorius and modified Sartorius scores: left and right axillae, left and right inframammary areas, intermammary area, left and right buttocks, left and right inguino-crural folds, perianal area, perineal area, and other. The presence of nodules, abscesses, fistulae, scars, and other findings will be recorded. The longest distance between two lesions and whether lesions are separated by normal skin will be recorded.

[7]IHS4 score (points) = (number of nodules multiplied by 1) + (number of abscesses multiplied by 2) + [number of draining tunnels (fistulae/sinuses) multiplied by 4]. A score of 3 or less signifies mild HS, a score of 4–10 signifies moderate HS and a score of 11 or higher signifies severe HS.

[8]"Patient-reported outcomes" (PROs) include the SF-36 v2, EQ-5D-5L, HiSQOL, DLQI, and WPAI-SHP questionnaires. The SF-36 v2 and EQ-5D-5L instruments are widely accepted global non-disease-specific tools to measure changes in subjects' health-related quality of life. The HiSQOL index (an HS-specific instrument) and DLQI instruments are designed to measure the impact of HS or skin disease on subjects' quality of life. The WPAI:SHP assesses the effect of general and specific health conditions on productivity losses.

[9] PK blood sample will be collected prior to the morning dose on Day 1 and at 0.5, 1, 2, and 3 (+/- 5 minutes) hours following dosing. Single PK samples will be collected pre-dosing at the subsequent visits as indicated. The date and time of the PK sample collection will be recorded. The date and time of the last dose of study drug prior to the PK sample collection will also be recorded. On study visit days when PK samples are collected, it is preferable that the subjects take the morning dose of study drug at the site following the collection of PK samples. PK will be performed in up to 150 subjects participating in the study. The Sponsor will continuously assess the number of subjects enrolled in the PK group. If at any point during enrollment, the number of PK subjects is below the required number, PK samples will be considered mandatory for the remaining subjects to be enrolled in the study. Additional consenting will be requested from subject for collection of PK samples.

[10] Subjects on placebo who complete the blinded, placebo-controlled 12-week period (Period 1) will be re-randomized 1:1 to receive 10 mg or 30 mg Avacopan twice daily during the 24-week active treatment period (Period 2).

[11] At Week 12, 24, and 36 visits, full drug accountability will be conducted on returned study drug.

[12] If consented, photography will be used for the purpose of publications and lesion assessment comparison. If feasible, it is recommended that photographs be taken at every visit.

[13] Health-economic information includes hospitalizations, emergency or urgent care visits, and lesion interventions due to HS. The WPAI:SHP questionnaire administration will follow the Patient-Reported Outcomes schedule.

[14] HS-PGA is an ordinal scale specific to HS that categorizes subjects into clear, minimal, mild, moderate, severe, or very severe disease.

[15] Biomarker plasma sample will be collected at Day 1, Week 12, and Week 36 or Early Termination visits in subjects who consented and at study sites adequately equipped to perform the collection. For Day 1, the biomarker sample collection must be done before the subject takes the first dose of study drug.

V. Interventions

A. Clinical Trial Material

Study subjects will receive active avacopan or placebo capsules as study drug. The study drug consists of hard gelatin capsules containing 10 mg avacopan or placebo administered orally. Avacopan and placebo bottles and capsules will be identical in appearance.

Subjects will be asked to take 3 capsules of study drug orally with water and preferably with food every morning, and 3 capsules with water and preferably with food in the evening approximately 12 hours after the morning dose, as instructed. Study drug will be taken for 36 weeks (252 days) continuously.

VI. General Analytical Considerations

A. Baseline and Study Day

Baseline
Baseline for analysis purposes is defined as the last assessment prior to treatment start date/time.

For Subject's Global Assessment of Skin Pain (Numerical Rating Scale, Skin Pain NRS), the average of the last 7 assessments prior to the date of first dose of study drug will be used as Baseline. If there are only 3 to 6 daily assessments available, the average of the available assessments will be used. If there are 2 or less assessments available, the subject's Baseline will be considered as missing.

Change from Baseline
Change from baseline is defined as the post baseline value minus the baseline value. Percent change from baseline is calculated as follows: Percent change = (Change from baseline / Baseline) * 100.

Reduction from Baseline
Reduction from baseline is defined as the baseline value minus the post baseline value. Percent reduction from baseline is calculated as follows: Percent reduction = (Reduction from baseline / Baseline) * 100.

Study Day
Study day for analysis purposes is defined as (date of event – treatment start date) (+ 1 if the event occurs after treatment start date).

B. Analysis Visit Window

All visits including scheduled, unscheduled and early termination visits will be assigned to analysis visits using analysis visit windows based on the actual date the assessment took place. The start day of the analysis window will be calculated as the midpoint between the scheduled assessment and previously scheduled assessment for that parameter. The end day of the analysis window will be calculated as the midpoint between the scheduled assessment and the next scheduled assessment for that parameter. Where multiple measurements for a particular parameter appear within an analysis window, the scheduled visit will be used. If no scheduled visit appears in the analysis window, the result closest to the target day will be used. If equidistant and both are unscheduled and/or early termination visits, the later result will be used for the summary measure.

Though all measures may not be used in data summaries (e.g., two lab measures within the same analysis visit window), all measurements appear in the datasets and listings. For subjects where the event date is missing, the study day and analysis window will also be missing. See Tables 1-3 below for the analysis windows.

Table 1 Mapping for Endpoints Measured at Study Day 1, Weeks 2, 4, 8, 12, 16, 20, 24, 28, 32, 36, and 44

Scheduled Visit	Target Study Day	Visit Window (Study Day)
Period 1: Study Day = visit date - date of first dose of study drug in Period 1 +1		
Baseline	1	≤ 1
Week 2	15	2 to 22
Week 4	29	23 to 43
Week 8	57	44 to 71
Week 12	85	72 to min (114, the study day of the randomization date for Period 2)
Period 2: Study Day = visit date - date of first dose of study drug in Period 2 +1		
Entry of Period 2	1	≤ 1
Week 16	29	2 to 43
Week 20	57	44 to 71
Week 24	85	72 to 99

Scheduled Visit	Target Study Day	Visit Window (Study Day)
Week 28	113	100 to 127
Week 32	141	128 to 155
Week 36	169	156 to 183
Week 44	197	≥ 184

Table 2 Mapping for Skin Pain NRS

Scheduled Visit	Target Study Day	Visit Window (Study Day)
Period 1: Study Day = visit date - date of first dose of study drug in Period 1 +1		
Baseline	-1	<1
Week 2	14	1 to 21
Week 4	28	22 to 42
Week 8	56	43 to 70
Week 12	84	71 to min (113, the study day of the randomization date for Period 2 -1)
Period 2: Study Day = visit date - date of first dose of study drug in Period 2 +1		
Entry of Period 2	-1	<1
Week 16	28	1 to 42
Week 20	56	43 to 70
Week 24	84	71 to 98
Week 28	112	99 to 126
Week 32	140	127 to 154
Week 36	168	155 to 182
Week 44	196	≥ 183

Table 3 Mapping for Endpoints Measured at Study Day 1, Weeks 4, 12, 16, 28, 36, and 44

Scheduled Visit	Target Study Day	Visit Window (Study Day)
Period 1: Study Day = visit date - date of first dose of study drug in Period 1 +1		
Baseline	1	≤ 1
Week 4	29	2 to 43
Week 12	85	44 to min (114, the study day of the randomization date for Period 2)
Period 2: Study Day = visit date - date of first dose of study drug in Period 2 +1		
Entry of Period 2	1	≤ 1
Week 16	29	2 to 43
Week 28	113	44 to 155
Week 36	169	156 to 197
Week 44	225	≥ 198

Data collected at all visits will be included in the data listings with visit presented as reported by the site.

C. Missing Data

Imputation for Efficacy variables:

The approaches listed below will be used for handling missing data:

- Non-responder imputation (NRI): Subjects who have missing data at the time point of interest are treated as though they did not respond to the treatment.
- Last Observation Carried Forward (LOCF): The LOCF analyses will use the last observed non-missing evaluation (last completed non-missing evaluation, from composite endpoint) from the previous visit within the particular period for efficacy measures assessed to impute missing data at later visits in the same period. Baseline efficacy evaluations will not be

carried forward. LOCF will be the primary approach in the analysis of continuous variable.

- Observed case (OC): Missing data are not imputed. Only subjects with available data at the given time point are considered.
- Multiple Imputation (MI) – Markov Chain Monte Carlo (MCMC) imputation: Using multiple imputation methodology, missing data in Period 1 are imputed based on the MCMC method, please see Section VIII D for details.

The following table depicts which missing data handling approaches should be used based on variable priority (primary, secondary, other) and variable type (binary, continuous).

Endpoints	Variable Type	Missing Data Handling Approach			
		NRI	MI	LOCF	OC
Primary Efficacy	Binary	P	SEN	SEN	
Primary Efficacy (COVID-19)	Binary	P		SEN	
Secondary/ Other Efficacy (selected ones only, see Page 47)	Binary	P	SEN		
	Continuous			P1	SEN

P=Primary method, SEN=Sensitivity method
1=For selected endpoints

Imputation for AE and Concomitant Medication Start Dates:

For analyses of AEs and concomitant medication usage, a complete date must be established in order to correctly identify the AE or medication as occurring during treatment or not. For purposes of imputing missing components of partially-reported start and stop dates for AEs and for medication use, the algorithms listed below will be followed. Start and stop dates of AEs or concomitant medication will be displayed as reported in the subject data listings (i.e., no imputed values will be displayed in data listings).

Partial AE and concomitant medication start dates will be imputed as follows:

- If year is missing, no imputation will be performed.
- If both day and month of the start date are missing, the day and month will be imputed with the day and month of the first study dose date if the year is equal to the year of first dose. Otherwise, the month and day is imputed as the first day of the year (01 Jan).

- If only day is missing, and if the year and month are equal to the first dose date, the day will be imputed as day of the first dose date. Otherwise, the day will be imputed as "01".

D. Multiple Study Centers

No adjustment for stratification by the study centers is planned. However, centers will be grouped by geographic region and a subgroup analysis will be conducted to determine if region impacts the primary endpoint.

E. Covariate Adjustment in Primary Analysis

Continuous variables of the secondary/other efficacy endpoints, change from Baseline to Week 12 in the Sartorius score, modified Sartorius score, IHS4 score, HS-PGA, inflammatory and non-inflammatory nodule count, abscess count, draining, non-draining and total fistula count, total AN count, SF-36 v2 (domains and component scores) and EQ-5D-5L (visual analogue scale and index), HiSQOL, DLQI, and WPAI:SHP will be analyzed using a mixed effects model for repeated measures. This model will include treatment group, visit, treatment-by-visit interaction, and randomization strata (Hurley Stage II or Hurley Stage III, anti-TNF drug use or anti-TNF drug naive) as factors, and corresponding baseline as covariate.

Change from Baseline to Week 12 in WPAI:SHP will be analyzed using an ANCOVA model which includes treatment group, randomization strata (Hurley Stage II or Hurley Stage III, anti-TNF drug use or anti-TNF drug naive) as factors, and corresponding baseline as covariate.

F. Sample Size Reassessment

No sample size reassessment is planned.

G. Interim Analyses or Timing of Analyses

No interim analyses planned.

H. Test Sizes

All confidence intervals will be two-sided with 95% coverage.

I. Multiple Comparisons/Multiplicity

For the primary efficacy endpoint testing, the avacopan 30 mg vs. placebo in HiSCR (H_{11}) and the avacopan 10 mg vs. placebo in HiSCR (H_{12}), a hierarchical procedure will be used to control the overall α at 0.05 level.

That is if H_{11} is rejected at a 2-sided alpha level of 0.05, then the H_{12} will be tested at 2-sided alpha level of 0.05.

For secondary efficacy endpoint testing, the avacopan 30 mg vs. placebo or the avacopan 10 mg vs. placebo will be tested separately with the $\alpha = 0.05$ level, provided that the corresponding primary endpoint testing of the same dose level is significant. As the two drug doses are two independent families, and the outcome of the study is to select one dose, not 2 doses. A fixed hierarchical testing sequence will be employed, as following:

H_2 : Change from Baseline in total AN count at Week 12

H_3 : NRS30 at Week 12

H_4 : % subjects with baseline Hurley Stage II who achieved AN count of 0, 1, or 2 at Week 12

H_5 : Reduction of IHS4 at Week 12

H_6 : Change from Baseline in inflammatory nodule count at Week 12

H_7 : Change from Baseline in abscess count at Week 12

H_8: Change from Baseline in draining fistula count at Week 12

H_9 : Change from Baseline in Modified Sartorius score at Week 12

Comparisons involving other efficacy endpoints and time points are considered supportive or exploratory and will be made at α=0.050 level (two-sided). No multiplicity adjustment will be made for these comparisons.

J. Analysis Populations

Analysis populations will be defined for use with various analyses.

Intent-to-Treat Population

The Intent-to-Treat (ITT) Population will include all subjects who are randomized and have received at least one dose of study drug, and who are not from study Site 138. Subjects from Site 138 will be excluded from the ITT Population due to potential misconduct and non-compliance at site 138 .

The ITT populations are defined for the two treatment periods as the following:

- The ITT1 Population in Period 1 is defined as all subjects who are randomized at baseline and have received at least one dose of study drug during Period 1.
- The ITT2 Population in Period 2 is defined as all subjects who have received at least one dose of study drug during Period 2.

Per Protocol Population

The per protocol (PP) population for period 1 will consist of ITT1 and do not have protocol deviations that could significantly affect the interpretation of the results for the primary endpoints. Subjects' inclusion/exclusion from the PP population will be determined and documented prior to the database lock and unblinding.

Per-Protocol Population in Period 1 (PP1) will include subjects in ITT1 who meet all the following criteria:

- Receive at least 70% of the planned study drug in Period 1 for subjects who complete Period 1 or receive at least one dose of study drug and discontinue Period 1
- Provide at least one post Baseline assessment on lesion count
- Meet the Hurly stage II: Total SCAR and Fistula counts > 0 at either screening or Day 1.
- Have Baseline AN count ≥ 3
- Have Baseline draining fistula count ≤ 20
- Do not take the following exclusionary medications during the exclusionary screening period or during Period 1:
- Any concomitant antibiotics for the treatment of HS (except the protocol allowed rescue medication which will result in counting the subjects as non-responders, or having the last observation prior to the use of these treatments carried forward).
- Anti-TNF-a (Humira), Methotrexate (MTX), cyclosporin, corticosteroids, and retinoids for any reason, or other medication for treatment of their HS that will confound the efficacy evaluation (to be determined and documented in the classification results prior to blind break).

Per-Protocol Population in Period 2 (PP2) will be defined separately.

Subjects included in the Per-Protocol Population will be analyzed as treated. That is, if a subject receives the treatment that is not the randomized assignment during their entire participation of a period, the subject will be analyzed according to the treatment that the subject actually received in the period.

Safety Population

The Safety Population will include all subjects who are randomized and have received at least one dose of study drug, excluding those subjects who are from study Site 138. Due to the reasons stated above safety data of subjects from site 138, including adverse events, will be listed separately only and not be included in the summary tables:

- The Safety Population in Period 1 (Safety1) is defined as all subjects who are randomized at baseline and have received at least one dose of study drug during Period 1.
- The Safety Population in Period 2 (Safety2) is defined as all subjects who received at least one dose of avacopan during Period 2.
- The All Avacopan Treated Population is defined as all subjects who receive at least one dose of avacopan in any treatment period.

The Safety Populations will be analyzed according to the assigned treatment group. In the event that a subject receives a treatment regimen that does not correspond to the assigned treatment consistently, the subject will be included in the analysis group of the treatment actually received.

Pharmacokinetic Population

Pharmacokinetic (PK) population is a subset of the Safety population and includes all subjects with at least one measurable plasma pharmacokinetic concentration of avacopan or its metabolite M1.

The PK populations are defined for the two treatment periods as the following:

- The PK1 Population in Period 1 is a subset of the Safety1 population and includes all subjects with at least one measurable plasma pharmacokinetic concentration of avacopan or its metabolite M1 in Period 1.
- The PK2 Population in Period 2 is a subset of the Safety2 population and includes all subjects with at least one measurable plasma pharmacokinetic concentration of avacopan or its metabolite M1 in Period 2.

Subjects with major protocol deviation including PK deviation or compliance issues may be excluded from the PK analyses (i.e. descriptive statistics) upon agreement with the Sponsor on case by case basis but all data will be listed. Values that are excluded from the analysis will be footnoted appropriately.

K. Subgroup Analysis

Please see section VIII D for subgroup analysis.

L. Data Display Characteristics

Data displays produced for this study will include three types: summary tables, data listings, and figures.

Data listings will simply list the data recorded on the case report form (CRF) or derived for each subject. They will be ordered by treatment, subject number, and

time of assessment. Additional levels of ordering may be employed as appropriate. Data listings will not display subject initials.

In general, summary tables for the double-blind treatment period (Period 1) will be presented by treatment group:

- Placebo
- Avacopan 10 mg
- Avacopan 30 mg
- Avacopan Total (i.e., Avacopan 10 mg + Avacopan 30 mg)

Summary tables for the 24-week active treatment period (Period 2) will be presented by the following treatment group:

- Placebo to Avacopan 10 mg
- Placebo to Avacopan 30 mg
- Avacopan 10 mg
- Avacopan 30 mg
- Avacopan Total (i.e., Placebo to Avacopan 10 mg + Placebo to Avacopan 30 mg + Avacopan 10 mg + Avacopan 30 mg)

Summary tables for the combined periods using the All Avacopan Treated Population will be presented by the following treatment group:

- Placebo to Avacopan 10 mg
- Placebo to Avacopan 30 mg
- Avacopan 10 mg
- Avacopan 30 mg
- Avacopan Total (i.e., Placebo to Avacopan 10 mg + Placebo to Avacopan 30 mg + Avacopan 10 mg + Avacopan 30 mg)

Continuous data will be summarized with the number of non-missing values, mean, standard deviation, minimum, median, and maximum. Categorical data will be summarized with the number of non-missing values and the numbers of values equal to each of the possible values. Percentages of subjects with each of the possible values will be calculated from the number of subjects in the relevant cohort of the corresponding analysis population, unless stated otherwise. Some continuous variables may also be grouped into categorical levels and evaluated in frequency tables. For efficacy endpoints, such as change and percent change from baseline, if the data is not normal distributed, the logarithmic transformation of the data will be performed prior to analysis, and geometric mean will be presented.

All statistical analyses will be performed using SAS® software, Version 9.3, or higher.

VII. Subject Accountability

A. Subject Disposition

The number of subjects who were screened, who screen failed (by reason), who were randomized, who completed Week 12, Week 36, and Week 44 of the study, respectively, and who withdrew early from the study, along with the reasons for withdrawal, will be presented by treatment group and period. A separate table for analysis populations will also be provided by treatment group.

Patient disposition data will be provided in a listing. A separate listing describing each subject's inclusion or exclusion status for each of the analysis sets will also be provided.

B. Protocol Deviations

Protocol deviations will be captured by the site and reviewed by medical monitor during the study. Classification between CSR reportable and non-CSR reportable protocol deviations will be decided by the study team prior to database lock. All protocol deviations will be listed and summarized by type, treatment group and period.

A listing of protocol deviations will be included.

C. Subject Characteristics

Demographic and Baseline Characteristics

Demographic and baseline information will be summarized for the ITT1 by treatment group:

- Age
- Sex (Male, Female)
- Ethnicity (Hispanic or Latino, Not Hispanic or Latino, Not Reported)
- Race (White, Black or African American, Native Hawaiian or Other Pacific Islander, Asian, American Indian or Alaskan Native, Other, Not Reported)
- Height (cm)
- Weight (kg)
- BMI (kg/m^2)
- Hurley Stage of HS
- HS disease duration (from time of diagnosis)
- HS lesion counts:
    - Number of AN counts
    - Number of inflammatory nodules

  - Number of abscesses
  - Number of draining fistulae
  - Number of AN count categories: <5, 5-10, ≥10
  - Number of hypertrophic scars
- Sartorius and modified Sartorius scores
- IHS4 score
- HS-PGA
- Anatomic location of HS
- Subject's global assessment of skin pain NRS
- Previous systemic HS treatment (includes TNF inhibitor use but excludes antibiotic use)
- Previous TNF inhibitor use
- Previous antibiotic treatment for HS
- Prior surgery for HS
- SF-36 v2
- EQ-5D-5L index
- HiSQOL index
- DLQI index
- Health Economics WPAI:SHP

A listing of demographic and baseline information will be provided.

The stratification factor values as collected in the electronic case report forms will be used for the summaries of baseline characteristics.

Medical History and Concurrent Procedures

The Medical Dictionary for Regulatory Activities (MedDRA; Version 21.1) will be used to code all medical history terms or concurrent procedures to a System Organ Class (SOC) and Preferred Term (PT). Medical History/concurrent procedures will be summarized by SOC in alphabetical order, preferred term in alphabetical order, and by treatment group. Subjects reporting more than one PT within a SOC will be counted only once for that SOC.

All medical history and concurrent procedures will be provided in a listing separately.

Prior and Concomitant Medications

Prior medications are defined as any medication taken prior to the first dose of blind study medication. Concomitant medications are defined as any medication taken on or after the first dose of blind study medication. A medication is classified as both prior and concomitant if it started prior to the first dose of blind study medication date and continued into the treatment period.

Prior/concomitant medications will be coded to therapeutic class and preferred term using the World Health Organization Drug Dictionary (WHO-DD), Version September 2018.

The number and percentage of subjects who had taken prior/concomitant medications will be summarized by therapeutic class and preferred term, and by treatment group. Subjects taking the same medication multiple times will only be counted once for that therapeutic class and preferred term. A subject level listing will also be presented.

Study Drug Exposure and Compliance

Study drug exposure will be determined using the morning/evening drug diary, study drug dispensing and return records. The study drug exposure including duration, total dose, average daily dose and compliance will be summarized for Period 1 and Period 2 separately for the safety population. Individual subject data listings will also be provided. If date of last dose is not available, the date of discontinuation from study will be used.

The following are the definitions of the drug exposure measures:

- Duration of Study Drug: Defined as Last Dose Date – First Dose Date +1.
- Total Dose: Defined as total dose of Avacopan taken in mg.
- Average Daily Dose (mg/day): Defined as [total dose of avacopan taken (mg) / treatment duration (days)] .
- Percent compliance

$$\text{Compliance (\%)} = 100 \times \frac{(\text{Total \# of capsules dispensed} - \text{Total \# of capsules returned})}{\text{Duration of Treatment} * 6}$$

Duration on treatment will be calculated using the last dose date – first dose date + 1. The first dose date in Period 1 and the last dose date in Period 2 will be collected from EDC, the last dose date in Period 1 and the first dose date in Period 2 will be determined from the drug accountability form (DA). If the last dose date in Period 1 is missing for subjects with early termination, then the date of last visit will be used as the last dose date in Period 1. For subjects who did not return data on DA, the date of last dose and compliance will be calculated based on subjects' daily dosing diary.

VIII. Efficacy Analyses

All efficacy analyses will be performed using the ITT population unless otherwise noted. All statistical testing will be two-sided, with the Type I error rate at $\alpha = 0.05$. Summary tables will be presented for the Period 1 and Period 2 separately. The stratification factor values as collected in the electronic case report forms will be used for all stratified efficacy analyses, and subgroup analyses.

A. Timing of Analyses

Primary Analysis: This study consists of a 12-week placebo-controlled, double-blind treatment period (Period 1), followed by a 24-week active treatment period (Period 2). When all subjects have completed or dropped out of Period 1, an unblinded analysis of the data collected from Period 1 can be conducted.

Follow-up Analysis: When all subjects have completed all of the Week 44 assessments or have dropped out of study, the follow-up analysis will be conducted on all data collected in the study.

B. Efficacy Endpoints

- **Primary efficacy endpoint**

Primary efficacy endpoint is the proportion of subjects achieving Hidradenitis Suppurativa Clinical Response (HiSCR) at Week 12. A response is defined as a reduction of at least 50% in abscess and inflammatory nodule count (AN) count with no increase in abscess count and no increase in draining fistula count compared to baseline.

$$\text{\% Reduction from Baseline in AN Count} = 100 \times \frac{(\text{Baseline AN Count} - \text{Week 12 AN Count})}{\text{Baseline AN Count}}$$

The proportion of subjects achieving HiSCR at Week 12 will be calculated as those subjects having % Reduction from Baseline in AN count ≥ 50 and (Baseline Abscess Count − Week 12 Abscess Count ≥ 0) and (Baseline Draining Fistula Count − Week 12 Draining Fistula Count) ≥ 0 at Week 12 divided by the total number of subjects in each treatment group.

The location and extent of HS will be assessed at visits in the Time and Events Table by recording the anatomic location(s) of the disease, as well as the number of HS inflammatory nodules, abscesses, fistulae, and scars in each of the locations in the "Anatomic Location and Number of HS Lesions" CRF. Twelve body areas will be evaluated: left and right axillae, left and right inframammary areas, intermammary area, left and right buttocks, left and right

inguino-crural folds, perianal area, perineal area, and other. Total number of nodules, abscesses, fistulae, and scars for each subject will be calculated respectively by sum of the number across the locations.

- **Secondary Efficacy Endpoints**

1) Change from Baseline in total AN count at Week 12.

   Change from Baseline = Week 12 Count – Baseline Count

2) Proportion of subjects achieving at least 30% reduction and at least 1 unit reduction from Baseline in the subject's global assessment of skin pain (NRS30) in subjects with a Baseline NRS of at least 3, evaluated at Week 12.

   Subjects will record the maximum severity pain on a numeric rating scale from 0 (no skin pain) to 10 (skin pain as bad as can be imagined) in a daily diary from one week prior to Day 1 through the Week 44 visit. Weekly averages of daily pain will be calculated based on subjects' daily diary recording of the worst pain experienced in the previous 24 hours. The weekly average score will be calculated based on the 7 available daily scores from the days within each visit window that are closest to the target day. If there are 3 to 6 assessments in the visit window, the average of the available assessments will be used. If more than one assessment is included on the same day, the worst assessment on that day will be chosen as the daily score. If there are more than 7 daily scores available and the 7th and 8th observations (ranked based on distance from the nominal day) are equidistant to the nominal day, the one after the nominal day will be used to calculate the weekly score. For any visit with less than 33 diary entries within the visit window, the weekly average will be recorded as missing.

   $$\%\ \text{Reduction from Baseline in NRS} = 100 \times \frac{(\text{Baseline NRS} - \text{Week 12 NRS})}{\text{Baseline NRS}}$$

   The proportion of subjects achieving NRS30 at Week 12 will be calculated as those subjects having % Reduction from Baseline in NRS ≥ 30 and (Baseline NRS – Week12 NRS ≥ 1) at Week 12 divided by the total number of subjects having Baseline NRS ≥ 3 in each treatment group.

3) Proportion of subjects with baseline Hurley Stage II who achieved an abscess and inflammatory nodule count of 0, 1, or 2 at Week 12.

   Number of abscess and inflammatory nodules is collected in the "Anatomic Location and Number of HS Lesions" CRF, please see the Primary Endpoint for details.

   The proportion of subjects with baseline Hurley Stage II achieving AN count of 0, 1, or 2 at Week 12 will be calculated as those subjects having baseline

Hurley Stage II and AN Count = 0, 1, or 2 at Week 12 divided by the total number of subjects having baseline Hurley Stage II in each treatment group.

4) Reduction and % reduction of IHS4 score relative to baseline at Week 12

The International Hidradenitis Suppurativa Severity Score (IHS4) is calculated as the following:

IHS4 score (points) = (number of nodules multiplied by 1) + (number of abscesses multiplied by 2) + [number of draining tunnels (fistulae/sinuses) multiplied by 4]

A score of 3 or less signifies mild HS, a score of 4 – 10 signifies moderate HS and a score of 11 or higher signifies severe HS.

$$\text{Reduction in IHS4} = \text{Baseline Score} - \text{Post Baseline Score}$$
$$\%\ \text{Reduction in IHS4} = 100 * \text{Reduction in IHS4/Baseline Score}$$

5) Change from Baseline in inflammatory nodule count at Week 12.

Change from Baseline = Week 12 Count – Baseline Count

6) Change from Baseline in abscess count at Week 12.

Change from Baseline = Week 12 Count – Baseline Count

7) Change from Baseline in draining fistula count at Week 12.

Change from Baseline = Week 12 Count – Baseline Count

8) Change from Baseline to Week 12 in the modified Sartorius score to quantify the severity change of HS.

Twelve body areas will be evaluated at visits specified in the Study Schedule to calculate the Sartorius **(in other efficacy endpoints)** and modified Sartorius [2009] scores:

- left and right axillae,
- left and right inframammary areas,
- intermammary area,
- left and right buttocks,
- left and right inguino-crural folds,
- perianal area and perineal area, and
- other (specify).

A score of 4 indicates the least severe disease, and higher scores indicate increasingly severe disease. There is no upper limit in the score.

The presence of nodules, abscesses, fistulae, scars, and other findings will be recorded in the EDC. The longest distance between two lesions and whether lesions are separated by normal skin will also be recorded. The final Sartorius and modified Sartorius [2009] scores will be the sum of each regional score.

Change from Baseline in Modified Sartorius Score = Week 12 Score – Baseline Score

The following are the algorithms on how to calculate the Sartorius and modified Sartorius [2009] scores for each regional score:

	Sartorius (2003)	Modified Sartorius [2009]
	For each anatomic region, calculate the regional score as follows:	For each anatomic region, sum up the following to obtain the subjects score:
1	3 points per region involved.	3 points per region involved.
2	Number and scores of lesions 2 points per nodule (or abscess) 4 points per fistula 1 point per scar 1 point per "other" lesion	Numbers and scores of lesions 1 point per nodule (or abscess) 6 point per fistula Note: Abscesses are considered the same as nodules since they are hard to distinguish from the other. Scars and other lesions are not considered.
3	The longest distance between two relevant lesions, or size if only one lesion < 5 cm, 2 points < 10 cm, 4 points > 10 cm, 8 points	The longest distance between two relevant lesions (or size of lesion if single) in each region < 5 cm = 1 point 5–10 cm = 3 points > 10 cm = 9 points
4	Are all lesions clearly separated by normal skin? In each region yes = 0; no = 6	Are all lesions clearly separated by normal skin? In each region yes = 0; no = 9

- **Other Efficacy Endpoints**

The following efficacy endpoints will be analyzed from Baseline to each timepoint up to Week 44, where applicable:

1) Change from Baseline in inflammatory nodule count, abscess count, draining fistula count, total AN count, and non-inflammatory nodule count, non-draining and total fistula count.

   Change from Baseline = Post Baseline Count – Baseline Count

2) % change from Baseline in inflammatory nodule count, abscess count, draining fistula count, total AN count, and non-inflammatory nodule count, non-draining and total fistula count among subjects who have at least one corresponding lesion at Day 1.

   % Change from Baseline = 100*Change from Baseline/Baseline Count

3) Proportion of subjects achieving HiSCR, $HiSCR_{75}$, and $HiSCR_{90}$.

   Please see the description of the Primary Endpoint for details.

$$\text{\% Reduction from Baseline in AN Count} = 100 \times \frac{(\text{Baseline AN Count} - \text{Post Baseline AN Count})}{\text{Baseline AN Count}}$$

   The proportion of subjects achieving HiSCR at Post Baseline visits will be calculated as those subjects having % Reduction from Baseline in AN count $\geq$ 50 and (Baseline Abscess Count – Post Baseline Abscess Count ≥ 0) and (Baseline Draining Fistula Count – Post Baseline Draining Fistula Count) ≥ 0

   at Post Baseline visits divided by the total number of subjects in each treatment group.

   $HiSCR_{75}$ or $HiSCR_{90}$ denotes 75% reduction or 90% reduction for the HiSCR definition instead of 50% reduction with no increase in the number of abscesses or draining fistulas count.

4) Proportion of subjects achieving at least 30% reduction and at least 1 unit reduction from Baseline in the subject's global assessment of skin pain (NRS30), in subjects with a Baseline NRS of at least 3.

   Please see the description of item 2) in Secondary Endpoints for details.

$$\text{\% Reduction from Baseline in NRS} = 100 \times \frac{\text{Baseline NRS} - \text{Post Baseline NRS}}{\text{Baseline NRS}}$$

   The proportion of subjects achieving NRS30 at Post Baseline visits will be calculated as those subjects having % Reduction from Baseline in NRS $\geq$ 30 and (Baseline NRS – Post Baseline NRS ≥ 1) at Post Baseline visits divided

by the total number of subjects having Baseline NRS ≥ 3 in each treatment group.

5) Proportion of subjects with baseline Hurley Stage II who achieved an abscess and inflammatory nodule count of 0, 1, or 2.

   Please see the description of item 4) in Secondary Endpoints for details.

   The proportion of subjects with baseline Hurley Stage II achieving AN count of 0, 1, or 2 at Post Baseline visits will be calculated as those subjects having baseline Hurley Stage II and AN Count = 0, 1, or 2 at Post Baseline visits divided by the total number of subjects having baseline Hurley Stage II in each treatment group.

6) Change from Baseline in the Sartorius score, modified Sartorius score, IHS4 score, and HS-PGA to quantify the severity change of HS.

   Sartorius score, modified Sartorius score, and IHS4 score has been described in items 1) and 3) of Secondary Endpoints.

   The Hidradenitis Suppurativa-Physician's Global Assessment (HS-PGA) is an ordinal scale specific to HS that categorizes subjects into clear, minimal, mild, moderate, severe, or very severe disease, and it was used successfully in a phase 2 interventional clinical trial. A recently developed six stage PGA was defined as follows:

Hidradenitis Suppurativa-Physician's Global Assessment		
Score	Short Descriptor	Definition
0	Clear	0 abscesses, 0 draining fistulas, 0 inflammatory nodules, and 0 non-inflammatory nodules
1	Minimal	0 abscesses, 0 draining fistulas, 0 inflammatory nodules, and only presence of non-inflammatory nodules
2	Mild	0 abscesses, 0 draining fistulas, and 1 to 4 inflammatory nodules or 1 abscess or draining fistula (sum of abscesses and draining fistulas is 1) and 0 inflammatory nodules
3	Moderate	0 abscesses, 0 draining fistulas, and ≥ 5 inflammatory nodules or 1 abscess or draining fistula and ≥1 inflammatory nodule or

		2 to 5 abscesses or draining fistulas (sum of abscesses and draining fistulas is 2 to 5) and < 10 inflammatory nodules
4	Severe	2 to 5 abscesses or draining fistulas and ≥10 inflammatory nodules
5	Very severe	>5 abscesses or draining fistulas (sum of abscesses and draining fistulas >5)

HS-PGA is collected from the "Hidradenitis Suppurativa-Physician's Global Assessment (HS-PGA)" CRF.

Change from Baseline in Score = Post Baseline Score – Baseline Score

7) Change from Baseline in subject-reported outcomes: SF-36 v2, EQ-5D-5L, HiSQOL and DLQI.

The SF-36v2, is the most widely used health-related quality-of-life measure in research to date. It measures the following 8 health domains as rated by the subjects over the past four weeks: Physical Functioning, Role Physical, Bodily Pain, General Health, Vitality, Social Functioning, Role Emotional, and Mental Health.

The SF-36 Physical and Mental Component Summary scores (PCS and MCS, respectively) are used to measure the two broad components, or aspects, of health-physical and mental. PCS and MCS are based on the aggregate of 8 health concepts described above and all of the 8 health domain scales are used to score both components summary measures.

One additional item asks respondents about health change over the past year.

The SF-36 will be used using Quality Metric's Health Outcomes™ Scoring Software. The software uses updates 2009 U.S. population norms and applies a Full Missing Score Estimation (Full MSE) method as follows:

- A health domain score (except the PF domain) will be estimated provided that at least one non-missing response is available within that domain
- For the PF domain item response theory will be used to develop a model for estimates of the missing score
- Regression methods are then applied to estimate the PCS and the MCS on the basis of the available domains.

The EQ-5D-5L (EuroQuality of Life-5 Domains-5 Levels) is a self-assessed, health related, quality of life questionnaire. The scale measures quality of life on a 5-component scale including mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each level is rated on scale that describes the degree of problems ranging from 1 to 5 in that area: no problems, slight problems, moderate problems, severe problems,

and extreme problems. EQ-5D-5L health states, defined by the EQ-5D-5L descriptive system, will be converted into a single index value.

This tool also has an overall health scale called visual analog scale where the rater selects a number from 0 (worst health state possible) to 100 (best health state possible).

The Hidradenitis Suppurativa Quality of Life (HiSQOL) is a simple, subject-administered, quality-of-life questionnaire that covers 9 domains including physical functions, activities of daily living, clothing choices due to HS, symptoms due to HS, emotional consequences, sexual functioning, social consequences, concentration consequences, and influences on work and study as assessed over the past 7 days. The HiSQOL index has been developed as an HS-specific instrument to measure the impact of HS on subjects' quality of life. Forms for these instruments will be completed by study subjects at visits specified in the Time and Events Table.

The scoring of each answer for the HiSQOL is as follows:

HiSQOL Scoring	
Response	Score
Extremely/ Unable to do due to my HS	4
Very much	3
Moderately	2
Slightly	1
Not at all/ I do not normally do this/ I am not sexually active/ I do not work or study	0
Question unanswered	0

The HiSQOL total score is calculated by summing the score of each question. Subscale scores are the total score for the items in each sub-domain. The higher the score, the more quality of life is impaired by hidradenitis suppurativa.

Dermatology Life Quality Index (DLQI),

The DLQI is a simple, subject-administered, 10-question, validated, quality-of-life questionnaire that covers 6 domains including symptoms and feelings, daily activities, leisure, work and school, personal relationships, and treatment, as assessed over the past week.

The scoring of each answer for the DLQI is as follows:

DLQI Scoring	
Response	Score
Very much	3
A lot	2
A little	1
Not at all	0
Not relevant	0
Question unanswered	0
Q7: 'prevented work or studying' = yes	3

Specifically:

- If one question is left unanswered this is scored 0 and the scores are summed and expressed as usual out of a maximum of 30.
- If two or more questions are left unanswered the questionnaire is not scored.
- If question 7 is answered 'yes' this is scored 3. If question 7 is answered 'no' but then either 'a lot' or 'a little' is ticked this is then scored 2 or 1. If "Not relevant" is ticked, the score for Question 7 is 0. If it is answered 'no', but the second half is left incomplete, the score will remain 0.
- If two or more response options are ticked, the response option with the highest score should be recorded.
- If there is a response between two tick boxes, the lower of the two score options should be recorded.

The DLQI is calculated by adding the score of each question. The maximum score is 30, and the minimum score is 0. The higher the score, the more quality of life is impaired.

Meaning of DLQI Scores

0-1 = no effect at all on subject's life

2-5 = small effect on subject's life

6-10 = moderate effect on subject's life

11-20 = very large effect on subject's life

21-30 = extremely large effect on subject's life

Change from Baseline in Score = Post Baseline Score – Baseline Score

8) Proportion of subjects who experienced flare, defined by at least a 25% increase in AN counts with a minimum increase of 2 AN counts relative to Baseline.

The AN count is the sum of all abscesses and inflammatory nodules counts across all regions.

$$\%\ \text{Increase in AN counts} = 100 \times \frac{(\text{AN Counts at Day x} - \text{AN Counts at Baseline})}{\text{AN Count at Baseline}}$$

The proportion of subjects who experienced flare at Post Baseline visits will be calculated as those subjects having % Increase in AN counts ≥ 25 and (Post Baseline AN Counts – Baseline AN Counts) ≥ 2 at Post Baseline visits divided by the total number of subjects in each treatment group.

9) Duration of flare in days (Observed case only)

Duration of flare in days is calculated from the day when flare is observed to the day prior to the observation that flare is no longer present; of note, there could be multiple periods that flares are observed, in which case, the total days from the multiple periods will be used.

10) Proportion of subjects who experience at least 25% increase in draining fistula counts with a minimum increase of 2 draining fistula counts relative to Baseline.

$$\%\ \text{Increase in Draining Fistula Counts} = 100 \times \frac{(\text{Post Baseline Draining Fistula Counts} - \text{Baseline Draining Fistula Counts})}{\text{Draining Fistula Count at Baseline}}$$

The proportion of subjects who experienced at least 25% increase in draining fistula counts with a minimum increase of 2 draining fistula counts relative to Baseline at Post Baseline visits will be calculated as those subjects having % Increase in Draining Fistula Counts ≥ 25 and (Post Baseline Draining Fistula Counts – Baseline Draining Fistula Counts) ≥ 2 at Post Baseline visits divided by the total number of subjects in each treatment group.

11) During Period 2, proportion of subjects with a loss of response, (LOR) defined as loss of at least 50% of AN count improvement achieved at Week 12 from Period 1.

$$\text{AN in Period 2} - \text{Week 12 AN} > \frac{1}{2}(\text{Baseline AN} - \text{Week 12 AN})$$

$$\text{i.e., AN in Period 2} > \frac{1}{2}(\text{Baseline AN} + \text{Week 12 AN})$$

Proportion of subjects with a LOR in Period 2 is calculated as those subjects having AN Counts in Period 2 $>$ 0.5 $\times$ (Baseline AN counts $-$Week 12 AN counts) divided by the total number of subjects in each treatment group in Period 2.

12) Time to LOR during Period 2.

For subjects with LOR, time to LOR (days) = date of LOR – date of the first dose of study drug in Period 2 + 1; For Subjects without LOR or lost follow-up, subjects will be censored at the date of the last assessment of number of HS lesions, i.e., time to LOR (days) = date of last assessment of number of HS lesions – date of the first dose of study drug in Period 2 + 1.

13) Proportion of subjects who received add-on antibiotic or other therapy for HS during Period 1 and Period 2 separately.

Add-on antibiotic or other therapies for HS are captured in the Concomitant Medication CRFs.

Proportion of subjects who received add-on antibiotic or other therapy will be calculated as those subjects who received add-on antibiotic or other therapy documented in the CM CRF post baseline divided by the total number of subjects in each treatment group in the safety population.

14) Proportion of subjects who start disallowed opioid pain therapy during Period 1 and Period 2 separately.

Oral and transdermal opioid analgesics (except for tramadol) for any reason are disallowed. Disallowed opioid pain therapy are captured in the Concomitant Medication (CM) CRFs.

Proportion of subjects who start disallowed opioid pain therapy will be calculated as those subjects who start disallowed opioid pain therapy documented in the CM CRF post baseline divided by the total number of subjects in each treatment group in safety population.

15) Proportion of subjects who undergo lesion intervention due to HS during Period 1 and Period 2 separately.

Lesion intervention due to HS is collected in AE and Concurrent Procedures CRFs.

Only two types of interventions are allowed: injection with intralesional triamcinolone acetonide suspension (intralesional Kenalog® rescue injections, i.e., triamcinolone acetonide, 10 mg total maximum per subject within a period no longer than 1 week) and incision and drainage. Any disallowed lesions interventions due to HS other than the two types described above will also be counted as "Lesion intervention due to HS" and will be recorded as major (=CSR reportable) protocol deviation.

During Period 1, a total of two protocol-allowed interventions are permissible. An intervention can occur on maximally two different lesions at the same visit or on the same lesion at two different study visits. The same lesion cannot be treated two times at the same visit. If a subject requires more than two interventions within the first 12 weeks, then they must discontinue taking blind study medication. However, these subjects will be requested to remain in the study and to complete all study procedures if possible.

During Period 2, maximally two interventions every 4 weeks are permitted. An intervention can occur on two different lesions at the same visit or on the same lesion at two different study visits. Within each 4-week period, the same type of intervention cannot be used two times on the same lesion. If a subject requires more than two interventions within a 4-week period or has two of the same interventions on the same lesion within that period, the subject must discontinue taking blind study medication. However, these subjects will be requested to remain in the study and to complete all study procedures if possible.

Proportion of subjects who undergo lesion intervention due to HS will be calculated as those subjects who have lesion intervention due to HS documented in the AE/Concurrent Procedure CRF post baseline divided by the total number of subjects in each treatment group in the safety population.

Proportion of subjects who received oral antibiotic rescue therapy will be calculated as those subjects who received oral antibiotic rescue therapy documented in the CM CRF post baseline divided by the total number of subjects in each treatment group in safety population.

16) Number of lesion interventions due to HS during Period 1 and Period 2 separately.

Lesion intervention due to HS is collected in Concurrent Procedures CRFs. Number of lesion interventions due to HS per subject will be calculated by period.

17) Health-economic information:

a. WPAI:SHP: Change from Baseline to each timepoint during Period 1 and Change from Week 12 to each timepoint during Period 2 and including the follow-up period;

Work Productivity and Activity Impairment Questionnaire (WPAI:SHP) assesses the effect of general and specific health conditions on productivity losses. Scores are calculated for four areas: percent work time missed due to HS, percent impairment while working due to HS, percent overall work impairment due to HS, and percent activity impairment due to HS. Forms for this instrument will be completed by study subjects at visits specified in the Time and Events Table to measure changes from baseline in work productivity and activity impairment.

b. Hospitalizations (cumulative) by period: Number of Hospitalizations total and due to HS, Number of days hospitalized, Number of days missed from work.

   Hospitalizations are captured in Heath Economics Information CRFs, cumulative number will be calculated by sum of the numbers from each visit.

c. Emergency or Urgent Care visits (cumulative) by period: Number of visits total and due to HS, Number of days (or hours where applicable) missed from work.

   Emergency or Urgent Care visits are captured in Heath Economics Information CRFs, cumulative number will be calculated by sum of the numbers from each visit.

d. Lesion interventions for HS by period: Number of days (or hours where applicable) missed from work due to HS lesion interventions.

   Lesion interventions for HS are captured in Heath Economics Information CRFs, cumulative number of days missed from work due to HS lesion interventions will be calculated by sum of the numbers from each visit.

C. General Approach

Lesions that received intervention (incision and drainage, or intralesional injection of corticosteroid) will be counted as permanently present from the date of the intervention.

In order to adjust for the impact of major protocol disallowed treatments on the results of the trial, efficacy assessments obtained after the start of major protocol disallowed treatments will be excluded from the analysis (unless otherwise specified). Subject will be counted as non-responders for the categorical variables and have their last observation carried forward for continuous variables.

Concomitant therapy that would qualify as a rescue medication will be reviewed prior to database lock.

Additional analyses may be performed for the primary efficacy variable adjusting for Baseline covariates in addition to Baseline Hurley Stage, if necessary. Results from any additional analyses will not be used as a substitute for the planned analyses, but may be used as supplemental information for the study report.

D. Primary Efficacy Endpoint Analysis

Primary Analysis

The primary efficacy objective is to evaluate the efficacy of avacopan compared to placebo in subjects with Hurley Stage II or III HS based on subjects achieving a Hidradenitis Suppurativa Clinical Response (HiSCR) after 12 weeks of treatment. To address this objective, the following hypotheses will be tested using the data collected from Period 1 after the last subject has completed Period 1:

- The null hypothesis (H_0) is that the avacopan group is not different from the placebo control group when comparing the HiSCR rate at Week 12.
- The alternative hypothesis (H_1) is that the avacopan group is different from the placebo control group when comparing the HiSCR rate at Week 12.

To address the multiplicity issue with two dose groups, a fixed hierarchical testing procedure will be used for the comparisons of the Avacopan 10 mg vs. Placebo and Avacopan 30 mg vs. Placebo to maintain the overall Type I error at 0.05. The hypothesis testing will test the Avacopan 30 mg vs. Placebo first: if the p-value is ≤ 0.05, then 10 mg Avacopan vs. Placebo will be tested at the $\alpha = 0.05$ level. If the test of Avacopan 30 mg vs. Placebo is not significant at the $\alpha = 0.05$ level, the 10 mg Avacopan vs. Placebo will be tested as an exploratory analysis.

To compare treatment groups, the proportion of subjects achieving HiSCR at Week 12 will be tested using the Cochran-Mantel-Haenszel (CMH) test, stratified by Hurley Stage (Stage II vs. III), concomitant antibiotic therapy with allowed antibiotics (Yes vs. No) and prior anti-TNF drug use (Treatment naïve vs. Previous treatment). The two-sided 95% confidence intervals for the difference in proportions (avacopan minus placebo control) will be calculated using the stratified Newcombe hybrid-score method. If the sample becomes sparse in one of the strata, the stratification factor will be removed from the stratified analysis.

The primary efficacy analysis will be carried out in the ITT1 Population. In addition, the primary efficacy endpoint will be analyzed in the PP1 Population.

The following provides sample code for implementing the CMH analysis for 30 mg avacopan vs. placebo :

```
ods output commonpdiff=pdiff1;
proc freq data=test(where=(trt in (0, 2)));
```

```
        table hs*antib*antitnf*trt*resp/cmh1 riskdiff (common
              cl=Newcombe) noprint;
run;
```

where

- HS is the variable for Hurley Stage (Stage II vs. III)
- ANTIB is the variable for allowed antibiotics (Yes vs. No)
- ANTITNF anti-TNF drug use (Treatment naive vs. Previous treatment)
- TRT is the variable for treatment (0 for placebo; 1 for avacopan 10 mg; 2 for avacopan 30 mg)
- RESP is the response variable for achieving HiSCR at Week 12 (0 for No, 1 for Yes).

Missing assessments were handled with the use of NRI as the primary approach to handling missing data.

Sensitivity Analyses

Several sensitivity analyses will be conducted on the primary efficacy endpoint in the ITT1 Population. Each are described below:

1) Imputation of Missing Values with Multiple Imputation (MI) Analysis

The Multiple Imputation analysis will be carried out in three steps;

- Imputation of missing data. The imputation will be generated for the three components of HiSCR (abscess, draining fistula, inflammatory nodule counts). The variables to be included in the imputation model are: Baseline Hurley Stage, Baseline antibiotic therapy, Baseline anti-TNF-α treatment, treatment group, and values at each visit from Baseline up to Week 12. For each HiSCR component, 100 'complete' data sets will be generated using SAS PROC MI utilizing the MCMC algorithm. The seeds for the imputation will be specified as the value "12061", "36102", and "94218" for imputation of abscess, draining fistula, and inflammatory nodule counts, respectively. The 'complete' data sets for the individual components of the HiSCR will be merged by imputation number to allow calculation of the HiSCR and creation of 100 'complete' data sets for evaluation of HiSCR. Subjects who received concomitant anti-TNF-α treatment or other treatments with clinically relevant impact on HS will be forced as non-responders for all of 100 'complete' data sets.

- Analysis of imputed data sets. A CMH test, stratified by Baseline Hurley Stage (II vs. III), Baseline antibiotic therapy (Yes vs. No), Baseline anti-TNF--α treatment (Treatment naïve vs. Previous treatment) will be used to analyze each 'complete' HiSCR data set.

- Synthesis of imputation and analysis results. SAS PROC MIANALYZE will be used to generate the final estimate of the proportion difference between avacopan and placebo groups and the corresponding 95% confidence interval.

2) Non-Responder Imputation for Any Add-on Antibiotics/Other Concomitant Medications Use

All subjects with any add-on antibiotics (any antibiotics other than the concomitant at Baseline) or with dose increase from baseline concomitant antibiotics prior to Week 12, regardless the reason for use; or other concomitant medications listed below will be imputed as non-responders:

1. Add-on antibiotics
2. Concomitant Anti-TNF-α treatment
3. Concomitant other systemic treatment for HS. E.g., methotrexate, cyclosporine, retinoids, and fumaric acid esters, etc.

3) LOCF for HiSCR

Subgroup Analysis

Subgroup analyses will be performed for the primary efficacy endpoint in the ITT1 Population.

The subgroups are defined as follows:

- Randomization stratification variables
  - Hurley Stage (II, III)
  - Concomitant antibiotic therapy (Yes, No)
  - Anti-TNF-α treatment (Treatment naïve, Previous treatment)
- Sex (Male, Female)
- BMI category: Normal (< 25 kg/m^2) , Overweight ($25 - < 30$ kg/m^2), obese ($30 - < 40$ kg/m^2), Morbid obesity (≥ 40 kg/m^2)
- Age at diagnosis of HS (< median, ≥ median)
- Age at study entry (<65 years vs. ≥ 65 years)
- Duration of HS (< median, ≥ median)
- Race (White vs. None-White)
- Baseline AN count (< 5, 5-10, >10)
- Baseline inflammatory nodule count (< median, ≥ median)
- Baseline abscess count (< median, ≥ median)

- Baseline draining fistula count (< median, ≥ median)
- Baseline modified Sartorius score (< median, ≥ median)
- Previous systemic HS treatment (includes TNF inhibitor use but excludes antibiotic use)
- Geographic distribution (Northeastern Region, Southeastern Region, Western Region, Midwestern Region, and Southwestern Region)

These summaries will be based on imputed data (NRI for binary variables). If the sample becomes sparse in one of the subgroup strata, the subgroup variable will be collapsed or removed from the subgroup analysis.

Exploratory Analysis

The avacopan 10 mg vs. 30 mg groups will be tested for the primary efficacy endpoint if the primary efficacy endpoint tests of both avacopan 10 mg vs placebo and avacopan 30 mg groups vs placebo are statistically significant in the ITT1 Population.

E. Secondary and Other Efficacy Endpoints Analyses

The following missing data handling rules will apply:

Secondary/Other Efficacy Endpoints

- For binary efficacy variables, NRI imputation method will be used for missing data. For NRS30 in Period 1, multiple imputation method will be applied similarly as the primary endpoint.
- For continuous efficacy variables, OC method will be performed.
- For % change from baseline in inflammatory nodule count, abscess count, non-inflammatory count, draining fistula, non-draining fistula, total fistula count, and total AN count, no statistical testing will be performed, only summary statistics will be provided.

For the secondary endpoints, multiplicity issue has been addressed in Section VI.I

1. Analysis of Secondary and Other Efficacy Endpoints in Period 1

The Secondary Efficacy Endpoints for Period 1 will be tested for the avacopan 30 mg vs. placebo, and avacopan 10 mg vs. placebo at Type I error at 0.05. The analyses will be carried out in the ITT1 and PP population. Analysis of other efficacy endpoints will be conducted from Baseline to each timepoint in the ITT1 population.

Categorical variables of change from Baseline to Week 12 will be analyzed using CMH test stratified for baseline Hurley Stage (Stage II vs. III), concomitant antibiotic therapy with allowed antibiotics (Yes vs. No) and anti-TNF-α drug use

(Treatment naïve vs. Previous treatment) similar to the analysis of the primary efficacy endpoint.

Continuous variables will be analyzed using a mixed effects model for repeated measures (MMRM). This model will include treatment group, visit, treatment-by-visit interaction, and randomization strata (Hurley Stage II or Hurley Stage III, previous anti-TNF-α drug use or anti-TNF-α drug naïve, concomitant antibiotic therapy use or not) as factors, and corresponding baseline as covariate. Subjects will be considered as repeated measure units over visits. A Toeplitz covariance matrix will be used to model the within-subject variance-covariance structure for the model errors. If the model does not converge using the unstructured (UN) covariance matrix, Toeplitz (TOEPH) covariance matrix will be used. If convergence is still not met, then AR(1) or compound symmetry (CS) will be used. Testing p-values, point estimates and corresponding 95% confidence intervals will be estimated for the difference between the avacopan group and the placebo control group across 12 weeks using linear contrast from the model.

The following provides sample code for implementing the MMRM analysis:

```
ods output lsmeans=lsmeans diffs=diffs;
proc mixed data=test;
class subjid trt visit antitnf hs;
model ch = trt visit trt*visit hs antib antitnf
           base/ddfm=kr cl alpha=0.05 solution;
repeated visit/ type=UN sub=subjid ;
lsmeans trt*visit/cl diff=control("0");
run;
quit;
```

where

- CH is the variable for change from baseline
- TRT is the variable for treatment (0 for placebo; 1 for avacopan 10 mg; 2 for avacopan 30 mg)
- HS is the Hurley Stage (II vs. III)
- VISIT is the variable for study visit
- ANTIB is the variable for concomitant antibiotic therapy with allowed antibiotics (0 for No; 1 for Yes)
- ANTITNF is the variable for anti-TNF drug use (Treatment naive vs. Previous treatment)
- BASE is the variable for the corresponding baseline value
- SUBJID is for subject

For the analysis of subject's global assessment of skin pain (NRS30) up to Week 12, subjects who received analgesics (to be determined by classification prior to database lock), for skin pain, or have undergone intervention will be counted as non-responder for categorical variables, and have their last pain assessment before the start of the analgesic or date of the incision and drainage (I&D) carried forward

for continuous variables, from the start day of the analgesics or day of the I&D until 14 days after the stop of analgesic use or day of the I&D.

LOCF and as observed analysis will be the sensitivity analyses for categorical variables and for continuous variables, respectively

Disallowed pain therapy will be determined by a blinded data review prior to database lock.

In addition, as a sensitivity analysis for Period 1, for selected secondary efficacy endpoints (continuous variables), an ANCOVA model with treatment, baseline value, Hurley Stage (Stage II vs. III), concomitant antibiotic therapy with allowed antibiotics (Yes vs. No) and anti-TNF drug use (Treatment naïve vs. Previous treatment) will be also applied based on OC; for selected secondary efficacy endpoints (binary variables), Multiple imputation (MI) method will be applied. The following are selected secondary efficacy endpoints:

- Reduction of IHS4 score relative to baseline over time (MMRM model, and ANCOVA model); Log ratio of IHS4 score to baseline over time (MMRM model).
- Proportion of subjects achieving at least 30% reduction and at least 1 unit reduction from Baseline in the subject's global assessment of skin pain (NRS30) in subjects with a Baseline NRS of at least 3, evaluated at each time point to Week 12 (NRI, and MI).
- Change from Baseline in the modified Sartorius score to quantify the severity change of HS (MMRM model, and ANCOVA model)
- The Short Form-36 version 2 (SF-36 v2) (MMRM model, and ANCOVA model)
- The EQ-5D-5L (EQ-5D-5L) (MMRM model, and ANCOVA model)
- The Hidradenitis Suppurativa Quality of Life (HiSQOL) Index (MMRM model, and ANCOVA model)
- the Dermatology Life Quality Index (DLQI) (MMRM model, and ANCOVA model)
- The Work Productivity and Activity Impairment Questionnaire: Specific Health Problem (WPAI:SHP), change from baseline to Week 12 (MMRM model, ANCOVA model)

For the other endpoints, only the primary analysis will be applied.

For the analysis of variables in Period 1, not by visit, such as: duration of flare, health-economic information on hospitalizations, emergency or urgent care visits,

and lesion interventions for HS, an analysis of covariance (ANCOVA) model with treatment, baseline value, Hurley Stage (Stage II vs. III), concomitant antibiotic therapy with allowed antibiotics (Yes vs. No), and anti-TNF drug use (Treatment naïve vs. Previous treatment) will be applied. Treatment comparisons between the avacopan group and the placebo control group will be presented.

2. **Analysis of Secondary and Other Efficacy Endpoints in Period 2**

The efficacy data for secondary and other endpoints by timepoint in Period 2 will be summarized using descriptive statistics in the ITT2 population.

For binary variables, Non-Responder Imputation will be used; for change from baseline variables, other than lesion counts endpoints, no missing data imputations will be performed.

For selected secondary efficacy endpoints specified above, a MMRM model including treatment group, visit, treatment-by-visit interaction, and randomization strata (Hurley Stage II or Hurley Stage III, previous anti-TNF-α drug use or anti-TNF-α drug naïve, concomitant antibiotic therapy use or not) as factors, and corresponding baseline as covariate will be applied in Period 2. Testing p-values, point estimates and corresponding 95% confidence intervals will be estimated for the difference the avacopan 30 mg vs. total placebo to avacopan group, and avacopan 10 mg vs. total placebo to avacopan group across visits in Period 2 using linear contrast from the model.

For the analysis of time to loss of response (LOR), the treatment difference between the avacopan 30 mg vs. placebo to avacopan 30 mg, avacopan 10 mg vs. placebo to avacopan 10 mg, avacopan 30 mg group vs. avacopan 10 mg group, and placebo to avacopan 30 mg vs. placebo to avacopan 10 mg will be analyzed using the stratified log-rank test with the baseline Hurley Stage (Stage II vs. III), concomitant antibiotic therapy with allowed antibiotics (Yes vs. No), and anti-TNF drug use (Treatment naïve vs. Previous treatment) as the stratification variables. Subjects without LOR or lost follow-up will be censored at the date of the last assessment of number of HS lesions. Kaplan-Meier curves will be provided by treatment group.

IX. Safety Analyses

Safety analyses will be performed in the Safety Analysis Set using the actual treatment group.

A. Adverse Events

All subjects will be assessed regularly for potential occurrence of adverse events (AEs) throughout the study until the subject is off study. The Medical Dictionary for Regulatory Activities (MedDRA; Version 21.1) will be used to code all adverse events to a System Organ Class (SOC) and Preferred Term (PT).

A treatment-emergent AE (TEAE) in Period 1 is defined as any AE that occurs after the first dose of study treatment in Period 1 and before the date of first dose of study treatment in Period 2.

A treatment-emergent AE (TEAE) in Period 2 is defined as any AE that occurs after the first dose of study treatment in Period 2.

A treatment-emergent AE (TEAE) for all avacopan, is defined as any AE that occurs after the first dose of avacopan treatment in study.

The incidence of TEAEs will be summarized and tabulated by treatment group and period.

The severity of each AE will be assessed as mild, moderate, severe, life-threatening, fatal, or death. Several occurrences of the same AE in one subject will be counted once and the one with maximum severity will be counted.

The relationship of each AE to the blind study medication will be classified as probably not related or possibly related. Several occurrences of the same AE in one subject will be counted once and the one with the maximum relationship to blind study medication will be counted. If an AE is missing the relationship to blind study medication, the event will be assumed to be probably related for analysis and summarization.

Adverse Events of Special Interest (AESI) include infections, hepatic enzyme elevations, neutropenia and lymphopenia, hypersensitivity/angioedema, creatine phosphokinase elevation, malignancies, and HS Lesion interventions.

Infections: The subject incidence of serious infections, severe infections (i.e., Grade 3+), and infections leading to subject withdrawal from the study will be summarized by treatment group. Infections will be identified by selecting terms with the Adverse Event System Organ Class of 'Infections and Infestations'.

Hepatic transaminase elevations:
The following lab abnormalities will be included in this category:

- Grade 3 or greater increased ALT or AST (>5 times the upper limit of normal)
- Grade 2 or greater increased ALT or AST (>3 times the upper limit of normal) with elevation of bilirubin to >2 times the upper limit of normal or INR >1.5

Neutropenia, lymphopenia and leukopenia:
The following lab abnormalities will be included in this category:

- Grade 3 or greater neutropenia (<1 x 109/L)

- Grade 4 lymphopenia (<0.2 x 109/L)
- Grade 3 or greater leukopenia (WBC count <2 x 109/L)

Hypersensitivity reactions: The subject incidence of adverse events associated with hypersensitivity/angioedema and urticaria will be summarized by treatment group. Terms will be based on the Standardized MedDRA Query for hypersensitivity.

Creatine Phosphokinase Elevation:

Grade 3 or greater CPK increase (>5 times the upper limit of normal)

Malignancies: Any malignancies will be reported as AEs.

HS Lesion Interventions: Lesion interventions due to HS will be recorded in CP.

The following summaries will be presented in tables:

- Overall summary of TEAEs
- TEAEs by SOC and PT. (The number and percentage of subjects with at least one TEAE, as classified by system organ class (SOC) and preferred term (PT), will be summarized by treatment received in the period of onset. For these summaries, subjects with multiple events will be counted only once per SOC and preferred term.)
- TEAEs by SOC, PT and maximum relationship to study drug
- TEAEs by SOC, PT and maximum Severity
- Serious TEAEs by SOC and PT
- TEAEs leading to withdrawal the study drug by SOC and PT
- TEAEs of special interest by SOC and PT
- TEAEs by PT in descending order
- SAEs by PT in descending order

The AE listings will be prepared, sorted chronologically within subjects for the following types of AEs. Each listing will include system organ class, preferred term, onset and end date, severity, relation to study drug, action taken, outcome, and SAE status.

- All AEs
- Serious AEs
- TEAEs leading to death
- TEAEs leading to withdrawal the study drug
- TEAEs of special interest

B. Clinical Laboratory Results

Laboratory test results include:

- Hematology: hemoglobin, hematocrit, RBC count, WBC count with differential, platelet count, mean cell hemoglobin, mean cell hemoglobin concentration, mean corpuscular volume.
- Serum Chemistry: liver panel (total bilirubin, lactate dehydrogenase, aspartate aminotransferase [AST], alanine aminotransferase [ALT]), renal panel (blood urea nitrogen, creatinine), creatine phosphokinase (CPK), albumin, sodium, potassium, magnesium, bicarbonate, chloride, calcium, inorganic phosphorus, glucose, total protein, alkaline phosphatase, total cholesterol, uric acid, serum amylase, and serum lipase.
- Urinalysis: nitrite, blood, and protein. If positive, microscopy will be performed;
- Virology (measured only at screening and may be measured at the local laboratory): hepatitis B surface antigen, hepatitis C antibodies, HIV 1 and 2 antibodies.
- TB screen: interferon γ release assay (IGRA).
- Pregnancy test: serum pregnancy test

All laboratory values will be converted to International System of Units (SI units). Lab results will be graded by the low/normal/high classifications based on normal ranges.

Laboratory data (actual values and change from baseline) will be summarized with the mean, standard deviation and median by treatment group and study visit. Laboratory data will be listed by treatment group, subject, and study visit. The CTCAE grades will be provided for selected Lab parameters. Abnormal laboratory values will be flagged.

The subject incidence of elevated laboratory values of ALT, AST, total bilirubin, alkaline phosphatase as assessed by the Central Laboratory, will be summarized by treatment group and Grade, as defined per Common Terminology Criteria for Adverse Events (CTCAE) Version 5. Shift from baseline to highest CTCAE grade during the study period will also be produced for these laboratory parameters based on the Central Laboratory measurements.

The same summary by treatment and CTCAE grade will be produced for subject incidence of elevated creatine phosphokinase (CPK), low neutrophils, low lymphocytes, low leukocytes, low hemoglobin, and low platelets. Shift from baseline to highest CTCAE grade during the study period will also be produced for these laboratory parameters. A by-subject listing including all data for these laboratory parameters including CTCAE grade for subjects with any abnormality will be provided

Urine and serum pregnancy test results for female only, HIV screening results, Hepatitis B/C screening results, and tuberculosis screening results will be listed.

C. Vital Signs

The following vital signs variables will be summarized: systolic blood pressure (mmHg), diastolic blood pressure (mmHg), temperature (°C) and heart rate (beats/min) and weight (kg). The following summaries will be provided:

- A summary of the absolute and change from Baseline value for each vital sign variable by treatment group, visit, and period.
- A summary of the number and percentage of subjects experiencing at least one Potentially Clinically Significant Vital Sign Findings by treatment group and visit.

Definitions of Potentially Clinically Significant Vital Sign Findings

Parameter (unit)	Category	Markedly Abnormal Values
Systolic Blood Pressure (mmHg)	Low Value	≤ 90 mmHg or decrease ≥ 20 mmHg from Baseline
	High Value	≥180 mmHg or increase ≥ 20 mmHg from Baseline
Diastolic Blood Pressure (mmHg)	Low Value	≤ 50 mmHg or decrease ≥ 15 mmHg from Baseline
	High Value	≥105 mmHg or increase ≥ 15 mmHg from Baseline
Heart Rate (bpm)	Low Value	≤ 50 bpm or decrease ≥ 15 bpm from Baseline
	High Value	≥ 120 bpm or increase ≥ 15 bpm from Baseline

A by-subject listing of all vital signs data will be provided. This listing will be presented by treatment group with abnormal vital sign values flagged as "L" or "H" accordingly.

D. 12-Lead Electrocardiogram (ECG)

ECG results will be listed in the data listings. Abnormal ECG findings will be listed by treatment group and study visit, and clinical significance of abnormalities indicated.

E. Physical Examination

Physical examination will include assessment of appearance/mental status, head, eyes, ears, nose, throat, and the following body systems: dermatologic, cardiovascular, respiratory, gastrointestinal, musculoskeletal and neurologic.

A listing will be provided for physical examination findings (i.e., normal or abnormal and whether each abnormality was clinically significant) by treatment, visit, body system.

F. Pharmacokinetic Analysis

Sample Collections

PK blood sample will be collected prior to the morning dose on Day 1 and at 0.5, 1, 2, and 3 (+/- 5 minutes) hours (per protocol 19 October 2018 and amendment 2.0) or at 0.5, 1, 2, 3, and 6 (+/- 5 minutes) hours (per protocol 19 October 2018 and amendment 1.0) following dosing. Single PK samples will be collected pre-dosing at the subsequent visits at Weeks 2, 4, 8, 12, 16, 20, 28, 36, and 44. The date and time of the PK sample collection will be recorded. The date and time of the last dose of study drug prior to the PK sample collection will also be recorded.

For plasma concentrations and PK parameters of avacopan (CCX168) or metabolite M1, descriptive statistics will include the number of values, arithmetic mean, standard deviation, and coefficient of variation (CV%), geometric mean (GM), CV% GM and 90% confidence intervals, minimum, median, and maximum.

Data Assembly

The concentration data of avacopan and metabolite M1 as reported by the bioanalytical laboratory will be used without rounding for all analysis.

Subjects with major protocol deviation such as PK deviation or compliance issues may be excluded from the PK analyses (i.e., descriptive statistics) upon agreement with the Sponsor on a case–by-case basis but all data will be listed. Values that are excluded from the analysis will be footnoted appropriately.

The following general rules will apply for handling missing data or concentration below the lower limit of quantification (BLQ):

- All BLQs will be set to "0" for Day 1 visit.
- If the BLQ occurs between 2 quantifiable drug concentrations within a full PK profile for Day 1 visit, the BLQ is excluded from all PK assessments.
- All BLQ will not be imputed for Post-Day 1 Visits and will be excluded from the descriptive summary.

- Missing pre-dose concentrations on Day 1 will be set as "0" and included in PK evaluation
- Missing post-dose values on Day 1 or missing trough values for post-Day 1 visits will not be imputed and will be excluded from PK evaluation
- Values that are excluded from the analysis will be documented appropriately.
- Pre-dose values on Day 1 that are greater than 5% of C_{max} will be documented. A sensitivity analysis (with inclusion and exclusion of the subjects with the pre-dose concentration >5% of C_{max}) will be presented and documented in the study report.

The sample time of the pre-dose samples on Day 1 will be uniformly considered as time "0".

For Day 1 visit, individual PK plots or individual PK parameters, since there is no collection window specified for post-dose PK sampling, will be based on actual times recorded. For concentration versus time descriptive statistical summaries and mean plot preparation, nominal time points will be used. If the difference of post-dose sampling time is > 5 minutes from the nominal sampling times for time points <2 hours or >5% for time points ≥2 hours, the corresponding concentration data will be excluded from the concentration summary and mean plot preparation, but will still be used in the individual plots and the calculation of PK parameter.

For Day 1 visit, if the pre-dose PK sample is collected after dosing, the corresponding concentration will be treated as a post-dose value and will be used for individual plotting and PK parameter calculation.

Post-Day 1 avacopan and its metabolite M1 plasma concentration results will be used to calculate trough plasma concentrations (C_{min}) over the course of the clinical trial.

For post-Day 1 visits (i.e. Weeks 2, 4, 8, 12, 16, 20, 28, 36, and 44), the allowable time window for PK sampling is ± 3 hours (i.e. 25% of dosing interval of 12 hours) for C_{min}. If the exact time (measured from dosing) is outside of the collection window, there is no dosing time, or there is a compliance issue, the corresponding concentration will be excluded from trough concentration versus time descriptive statistical summaries and median plot preparation, but will still be used in the individual plots.

Non-numerical readouts except BLQ [e.g., NA (not available) or NRE (not reliable)] will be set as missing and will not be included in the PK evaluation.

Pharmacokinetic Concentration

For subjects with serial samples collected on Day 1:

Individual plasma concentrations of avacopan and its metabolite M1 will be listed, plotted, and summarized descriptively and graphically by treatment group.

- Individual plasma concentrations of avacopan and its metabolite M1 will be plotted on a linear and semi-logarithmic scale against actual sampling time points;
- Concentration data will be summarized by treatment at each nominal time point descriptively;
- Mean (±SD) plasma concentrations of avacopan and its metabolite M1 will be plotted on a linear and semi-logarithmic scale against nominal time points by treatment.

For subjects with pre-dose sample collected on Post-Day 1 visits:

- Individual plasma concentrations of avacopan and its metabolite M1 will be plotted by treatment on a linear scale against visits;
- Trough concentration data, within allowable time windows, will be summarized by treatment at each visit descriptively.
- In addition, median trough concentrations (C_{min}) of avacopan and its metabolite M1 will be plotted overlaid with the scatter plot of individual trough concentrations by treatment on a linear scale against visits.
  For subjects who received active treatment in Period 1, this plot will include visits in both Period 1 and Period 2.
- The average steady state trough concentration of avacopan or its metabolite M1 for each individual subject will be calculated for a to-be-defined steady-state time period (each subject needs to have at least 3 time points in this steady state period for the calculation). The global average steady state trough concentration of avacopan or its metabolite M1 will be listed and summarized descriptively.
- For subjects who received placebo in Period 1 (Week 1 through Week 12), the Period 2 concentration data (Week 16 through Week 36) will be analyzed separately from the subjects who received active treatment in Period 1.

<u>Pharmacokinetic Parameter</u>

PK analysis will be performed in the PK population. The following parameters for avacopan (CCX168) or metabolite M1 will be determined, where possible:

- C_{max} Maximum plasma concentration
- T_{max} Time of maximum plasma concentration
- AUC_{0-3h} Area under the plasma concentration-time curve from Time 0 to Hour 3 on Day 1
- AUC_{0-6h} Area under the plasma concentration-time curve from Time 0 to Hour 6 on Day 1
- C_{min} Trough level plasma concentrations at post-Day 1 visits

PK parameter calculations will be based on continuous time points from Day 1 (including Time 0).

Plasma PK parameters will be calculated by standard non-compartmental analysis. The actual collection times will be used for PK parameter calculation. The linear trapezoidal rule method (equivalent to the Linear Trapezoidal Linear Interpolation in WinNonlin® Professional) will be used in the computation of AUCs.

PK parameters will be listed and summarized by treatment descriptively.

The derived PK parameters (C_{max}, AUC, T_{max} and C_{min}) will be rounded to three significant digits.

The relationship between PK parameters and efficacy endpoints such as HiSCR may also be evaluated.

G. Biomarker Assessments

A summary of the absolute and change from Baseline value for each available biomarker variable will be presented by treatment group, visit, and period.

A listing will be provided for subjects who have biomarker assessments.

H. Photographic Assessments

A listing will be provided for subjects who received photography by visit.

I. Supplementary Analyses Related to COVID-19

In order to describe the impact of COVID-19 on the current study, the following will be summarized in the tables/listings:

- Subjects discontinued from the study due to COVID-19 related on the study exit form.
- Add a new table for protocol deviations due to COVID-19.
- Add information for TEAE (COVID-19) in Overall AE summary table.
- Subjects with study visits missed or altered (including home health visit, phone/video conferencing, other) due to COVID-19. This may include altered or cancelled study visits and data collections because of a temporary study site closure during certain time periods of the COVID-19 pandemic.
- An additional sensitivity analysis of the primary endpoint will be performed in the ITT population based on the LOCF and NRI imputation method.
    - LOCF: For any week 12 assessments performed via altered study visits or missing due to COVID019, last prior unaffected lesion assessment will be carried forward to week 12,

- NRI: impute any week 12 assessments performed via altered study visits or missing due to COVID019 as non-responders.

X. Changes from the Protocol Planned Analyses

The following changes from the protocol planned analyses are noted in this plan:

Changes	Protocol	SAP
Updated Adverse Events of Special Interest (AESI) definition	Page 63: 7.2.3.3. Adverse Events of Special Interest (AESI) The following findings, as defined below, must be reported as AEs and will be considered adverse events of interest (AESI): Infections: For medically important infections, the organisms involved in the infection need to be determined whenever possible and be documented in the EDC. All local and national vaccination recommendations should be followed. Hepatic Transaminase Elevation: • Grade 3 or greater increased ALT or AST (>5 times the upper limit of normal) • Grade 2 or greater increased ALT or AST (>3 times the upper limit of normal) with elevation of bilirubin to >2 times the upper limit of normal or INR >1.5 Neutropenia, Lymphopenia, and Leuokopenia: • Grade 3 or greater neutropenia (<1 x 109/L) • Grade 4 lymphopenia (<0.2 x 109/L) • Grade 3 or greater leukopenia (WBC count <2 x 109/L) Creatine Phosphokinase Elevation: • Grade 3 or greater CPK increase (>5 times the upper limit of dosing) Hypersensitivity Reactions: • urticaria • angioedema Malignancies: • Any malignancies will be reported as AEs. HS Lesion Interventions: • Lesion interventions as described in Section 5.7 will be recorded as AE.	Adverse Events of Special Interest (AESI) include infections, hepatic enzyme elevations, neutropenia and lymphopenia, hypersensitivity/angioedema, creatine phosphokinase elevation, malignancies, and HS Lesion interventions. **Infections:** The subject incidence of infections, serious infections, severe infections (i.e., Grade 3+), and infections leading to subject withdrawal from the study will be summarized by treatment group. Infections will be identified by selecting terms with the Adverse Event System Organ Class of 'Infections and Infestations'. **Hepatic transaminase elevations**: The following lab abnormalities will be included in this category: • Grade 3 or greater increased ALT or AST (>5 times the upper limit of normal) • Grade 2 or greater increased ALT or AST (>3 times the upper limit of normal) with elevation of bilirubin to >2 times the upper limit of normal or INR >1.5 **Neutropenia, lymphopenia and leukopenia:** The following lab abnormalities will be included in this category: • Grade 3 or greater neutropenia (<1 x 109/L) • Grade 4 lymphopenia (<0.2 x 109/L) • Grade 3 or greater leukopenia (WBC count <2 x 109/L)

		Hypersensitivity reactions: The subject incidence of adverse events associated with hypersensitivity/angioedema and urticaria will be summarized by treatment group. Terms will be based on the Standardized MedDRA Query for hypersensitivity. **Creatine Phosphokinase Elevation**: Grade 3 or greater CPK increase (>5 times the upper limit of normal) **Malignancies:** Any malignancies will be reported as AEs. **HS Lesion Interventions:** Lesion interventions due to HS will be recorded in CP.
The handling of multiplicity issues has been modified	Page 67: To address the multiplicity issue with two dose groups, a fixed hierarchical testing sequence will be employed. The hypothesis testing will test the avacopan 30 mg vs. placebo first. If the test is significant at the $\alpha = 0.05$ level, the avacopan 10 mg vs. placebo will be tested at the $\alpha = 0.05$ level. If the test of avacopan 30 mg vs. placebo is not significant at the $\alpha = 0.05$ level, the avacopan 10 mg vs. placebo will be tested as an exploratory analysis.	For the primary efficacy endpoint testing, the avacopan 30 mg vs. placebo in HiSCR (H_{11}) and the avacopan 10 mg vs. placebo in HiSCR (H_{12}), a hierarchical procedure will be used to control the overall α at 0.05 level. That is if H_{11} is rejected at a 2-sided alpha level of 0.05, then the H_{12} will be tested at 2-sided alpha level of 0.05. For secondary efficacy endpoint testing, the avacopan 30 mg vs. placebo or the avacopan 10 mg vs. placebo will be tested separately with the $\alpha = 0.05$ level, provided that the corresponding primary endpoint testing of the same dose level is significant. As the two drug doses are two independent families. And the outcome of the study is to select one dose, not 2 doses. A fixed hierarchical testing sequence will be employed.
The primary analysis population's name has been changed from mITT to ITT	Page 68: Modified Intent-to-Treat (mITT) Population	Intent-to-Treat (ITT) Population
Definitions of ITT, Safety and Per Protocol Populations.	Page 68:	**Intent-to-Treat Population** The Intent-to-Treat (ITT) Population will include all subjects who are

	8.4.1 Modified Intent-to-Treat Population For the purposes of efficacy data analysis, the Modified Intent-to-Treat (mITT) population will be used. This population will include all subjects who are randomized and received at least one dose of study drug. The efficacy population will be analyzed according to the treatment group each subject is randomized to. The mITT population will be the primary analysis population for the efficacy analysis. The mITT populations are defined for the two treatment periods as the following: • The mITT1 Population in Period 1 is defined as all subjects who are randomized at baseline and have received at least one dose of study drug during Period 1. • The mITT2 Population in Period 2 is defined as all subjects who have received at least one dose of study drug during Period 2. **8.4.2 Per-Protocol Population** The Per-Protocol (PP) population will consist of all randomized subjects who receive at least one dose of study drug and do not have protocol deviations that could significantly affect the interpretation of the results for the primary endpoints. Subjects' inclusion/exclusion from the PP population will be determined and documented prior to the database lock and unblinding. **8.4.3 Safety Population**	randomized and have received at least one dose of study drug, and who are not from study Site 138. Subjects from Site 138 will be excluded from the ITT Population due to potential misconduct and non-compliance at site 138 . The ITT populations are defined for the two treatment periods as the following: • The ITT1 Population in Period 1 is defined as all subjects who are randomized at baseline and have received at least one dose of study drug during Period 1. • The ITT2 Population in Period 2 is defined as all subjects who have received at least one dose of study drug during Period 2. **Per Protocol Population** The per protocol (PP) population for period 1 will consist of ITT1 and do not have protocol deviations that could significantly affect the interpretation of the results for the primary endpoints. Subjects' inclusion/exclusion from the PP population will be determined and documented prior to the database lock and unblinding. Per-Protocol Population in Period 1 (PP1) will include subjects in ITT1 who meet all the following criteria: – Receive at least 70% of the planned study drug in Period 1 for subjects who complete Period 1 or receive at least one dose of study drug and discontinue Period 1 – Provide at least one post Baseline assessment on lesion count – Meet the Hurly stage II: Total SCAR and Fistula counts > 0 at either screening or Day 1. – Have Baseline AN count ≥ 3 – Have Baseline draining fistula count ≤ 20

	The Safety Populations are defined as the following: • The Safety Population in Period 1 is defined as all subjects who are randomized at baseline and have received at least one dose of study drug during Period 1. • The Safety Population in Period 2 is defined as all subjects who received at least one dose of avacopan during Period 2. • The All Avacopan Treated Population is defined as all subjects who receive at least one dose of avacopan in any treatment period. The Safety Populations will be analyzed according to the assigned treatment group. In the event that a subject receives a treatment regimen that does not correspond to the assigned treatment, the subject will be included in the analysis group of the treatment actually received.	– Do not take the following exclusionary medications during the exclusionary screening period or during Period 1: – Any concomitant antibiotics for the treatment of HS (except the protocol allowed rescue medication which will result in counting the subjects as non-responders, or having the last observation prior to the use of these treatments carried forward). – Anti-TNF-a (Humira), Methotrexate (MTX), cyclosporin, corticosteroids, and retinoids for any reason, or other medication for treatment of their HS that will confound the efficacy evaluation (to be determined and documented in the classification results prior to blind break). Per-Protocol Population in Period 2 (PP2) will be defined separately. Subjects included in the Per-Protocol Population will be analyzed as treated. That is, if a subject receives the treatment that is not the randomized assignment during their entire participation of a period, the subject will be analyzed according to the treatment that the subject actually received in the period. **Safety Population** The Safety Population will include all subjects who are randomized and have received at least one dose of study drug, excluding those subjects who are from study Site 138. Due to the reasons stated above safety data of subjects from site 138, including adverse events, will be listed separately only and not be included in the summary tables: • The Safety Population in Period 1 (Safety1) is defined as all subjects who are randomized at

		baseline and have received at least one dose of study drug during Period 1. • The Safety Population in Period 2 (Safety2) is defined as all subjects who received at least one dose of avacopan during Period 2. • The All Avacopan Treated Population is defined as all subjects who receive at least one dose of avacopan in any treatment period. The Safety Populations will be analyzed according to the assigned treatment group. In the event that a subject receives a treatment regimen that does not correspond to the assigned treatment consistently, the subject will be included in the analysis group of the treatment actually received.
Adding Ranked Secondary Endpoints	Page 69: • Secondary Efficacy Endpoints Secondary efficacy endpoints include: 1. Reduction of IHS4 score relative to baseline at Week 12; 2. Proportion of subjects achieving at least 30% reduction and at least 1 unit reduction from Day 1 in the subject's global assessment of skin pain (NRS30) in subjects with a baseline NRS of at least 3, evaluated at Week 12; weekly averages of daily pain will be calculated based on subjects' daily diary recording of the worst pain experienced in the previous 24 hours; 3. Change from Day 1 to Week 12 in the modified Sartorius score to quantify the severity change of HS;	The following secondary efficacy endpoints will be analyzed according to the rank order as the follows: 1. Change from Baseline in total AN count at Week 12; 2. Proportion of subjects achieving at least 30% reduction and at least 1 unit reduction from Baseline in the subject's global assessment of skin pain (NRS30) in subjects with a Baseline NRS of at least 3, evaluated at Week 12; weekly averages of daily pain will be calculated based on subjects' daily diary recording of the worst pain experienced in the previous 24 hours; 3. Proportion of subjects with baseline Hurley Stage II who achieved an abscess and

	4. Proportion of subjects with baseline Hurley Stage II who achieved an abscess and inflammatory nodule count of 0, 1, or 2 at Week 12.	inflammatory nodule count of 0, 1, or 2 at Week 12; 4. Reduction of IHS4 score relative to baseline at Week 12; 5. Change from Baseline in inflammatory nodule count at Week 12; 6. Change from Baseline in abscess count at Week 12; 7. Change from Baseline in draining fistula count at Week 12; 8. Change from Baseline to Week 12 in the modified Sartorius score to quantify the severity change of HS.
Added other efficacy endpoints	Page 69 • Other Efficacy Endpoints The following efficacy endpoints will be analyzed from Day 1 to each timepoint up to Week 44, where applicable: 1. Proportion of subjects achieving HiSCR; 2. Proportion of subjects achieving at least 30% reduction and at least 1 unit reduction from Day 1 in subject's global assessment of skin pain (NRS30), in subjects with a Day 1 NRS of at least 3; 3. Proportion of subjects with baseline Hurley Stage II who achieved an abscess and inflammatory nodule count of 0, 1, or 2; 4. Change from Day 1 in inflammatory nodule count, abscess count, draining fistula count, and total AN count;	• Other Efficacy Endpoints The following efficacy endpoints will be analyzed from Baseline to each timepoint up to Week 44, where applicable: 1. Proportion of subjects achieving HiSCR, $HiSCR_{75}$, and $HiSCR_{90}$; 2. Proportion of subjects achieving at least 30% reduction and at least 1 unit reduction from Baseline in the subject's global assessment of skin pain (NRS30), in subjects with a Baseline NRS of at least 3; 3. Proportion of subjects with baseline Hurley Stage II who achieved an abscess and inflammatory nodule count of 0, 1, or 2; 4. Change from Baseline in inflammatory nodule count,

	5. Change from Day 1 in the Sartorius score, modified Sartorius score, IHS4 score, and HS-PGA to quantify the severity change of HS; 6. Change from Day 1 in patient-reported outcomes: SF-36 v2, EQ-5D-5L, HiSQOL, and DLQI; 7. Proportion of subjects who experienced flare, defined as an at least 25% increase in AN counts with a minimum increase of 2 AN lesions relative to Day 1; 8. Duration of flare in days (calculated from the day when flare is observed to the day prior to the observation that flare is no longer present; of note, there could be multiple periods that flares are observed, in which case, the total days from the multiple periods will be used); 9. Proportion of subjects who experience at least 25% increase in draining fistula counts with a minimum increase of 2 draining fistula counts relative to Day 1; 10. During Period 2, Proportion of subjects with a loss of response (LOR), defined as loss of at least 50% of AN count improvement achieved from Day 1 to Week 12; 11. Time to LOR during Period 2; 12. Proportion of subjects who received oral antibiotic rescue therapy; 13. Proportion of subjects who start disallowed opioid pain therapy; 14. Proportion of subjects who undergo lesion intervention due to HS;	abscess count, non-inflammatory count, draining fistula, non-draining fistula, total fistula count, and total AN count; 5. % Change from Baseline in inflammatory nodule count, abscess count, non-inflammatory count, draining fistula, non-draining fistula, total fistula count, and total AN count among subjects who have at least one corresponding lesion at Baseline. 6. Change from Baseline in the Sartorius score, modified Sartorius score, IHS4 score, and HS-PGA to quantify the severity change of HS; 7. Change from Baseline in subject-reported outcomes: SF-36 v2, EQ-5D-5L, HiSQOL and DLQI; 8. Proportion of subjects who experienced flare, defined by at least a 25% increase in AN counts with a minimum increase of 2 AN lesions relative to Baseline; 9. Duration of flare in days (calculated from the day when flare is observed to the day prior to the observation that flare is no longer present; of note, there could be multiple periods that flares are observed, in which case, the total days from the multiple periods will be used); 10. Proportion of subjects who experience at least 25%

	15. Number of lesion interventions due to HS; 16. Health-economic information: a. WPAI:SHP: Change from Day 1 to each timepoint during Period 1 and Change from Week 12 to each timepoint during Period 2 and including the follow-up period; b. Hospitalizations (cumulative): Number of Hospitalizations total and due to HS, Number of days hospitalized, Number of days missed from work; c. Emergency or Urgent Care visits (cumulative): Number of visits total and due to HS, Number of days (or hours where applicable) missed from work; d. Lesion interventions for HS (cumulative): Number of days (or hours where applicable) missed from work due to HS lesion interventions.	increase in draining fistula counts with a minimum increase of 2 draining fistula counts relative to Baseline; 11. During Period 2, proportion of subjects with a loss of response, (LOR) defined as loss of at least 50% of AN count improvement achieved from Week 12 in Period 1; 12. Time to LOR during Period 2; 13. Proportion of subjects who received oral antibiotic rescue therapy during Period 1 and Period 2 including the follow-up period separately; 14. Proportion of subjects who start disallowed opioid pain therapy during Period 1 and Period 2 including the follow-up period separately; 15. Proportion of subjects who undergo lesion intervention due to HS during Period 1 and Period 2 including the follow-up period separately; 16. Number of lesion interventions due to HS during Period 1 and Period 2 including the follow-up period separately; 17. Health-economic information: a. WPAI:SHP: Change from Baseline to each timepoint during Period 1 and Change from Week 12 to each timepoint during Period 2 and including the follow-up period; b. Hospitalizations (cumulative): Number of Hospitalizations total and due to HS, Number of days

		hospitalized, Number of days missed from work during Period 1 and Period 2 including the follow-up period separately; c. Emergency or Urgent Care visits (cumulative): Number of visits total and due to HS, Number of days (or hours where applicable) missed from work during Period 1 and Period 2 including the follow-up period separately; d. Lesion interventions for HS: Number of days (or hours where applicable) missed from work due to HS lesion interventions during Period 1 and Period 2 including the follow-up period separately.
Updated the definition of non-responders	Page 71: The following are examples for subjects who will be counted as non-responder: • Concomitant anti-TNF-α treatment or other treatments with clinically relevant impact on HS; • Antibiotic rescue therapy; • Non-protocol specified lesion intervention, or higher than permitted frequency or number of pre-specified lesion interventions. Protocol-specified lesion interventions will not be counted as Non-responder.	Lesions that received intervention (incision and drainage, or intralesional injection of corticosteroid) will be counted as permanently present from the date of the intervention. In order to adjust for the impact of major protocol disallowed treatments on the results of the trial, efficacy assessments obtained after the start of major protocol disallowed treatments will be excluded from the analysis. Subject will be counted as non-responders for the categorical variables and have their last observation carried forward for continuous variables.

Missing data imputation method has been updated	Page 71: Missing Data Analysis: For missing data analysis, the following imputation rules will be applied to categorical and continuous efficacy variables, respectively: • Categorical Efficacy Variables: NRI will be the primary approach for categorical variables. The NRI analysis will count subjects who have missing values at a specific visit as non-responders for that visit. LOCF will be the secondary approach for the missing data analysis. • Continuous Efficacy Variables: LOCF will be the primary approach for continuous efficacy variables of missing data. The LOCF analyses will have their valid efficacy assessments from the previous visits to impute missing data at later visits. Baseline efficacy evaluations will not be carried forward.	1) Added MI method for selected binary secondary efficacy endpoints. 2) Added ANCOVA model for selected binary secondary efficacy endpoints.
Added CTCAE grades for lab parameters	Page 73: No CTCAE was mentioned in the protocol.	Added analyses by CTCAE grades.
Added more subgroup analyses.	Page 75: The analysis of the efficacy endpoints may be adjusted by the following variables in the form of covariate analysis, stratified analysis, and/or subgroup analysis: • Randomization stratification variables - Hurley Stage II vs. Hurley Stage III - Concomitant antibiotic therapy (Yes vs. No) - Anti-TNF-α treatment (Treatment naïve vs. Previous treatment) • Sex • BMI • Baseline weight • Age at diagnosis of HS • Age at study entry • Duration of HS • Subject's age, race, and ethnicity • Baseline inflammatory nodule count • Baseline abscess count • Baseline draining fistula count • Baseline Sartorius and modified Sartorius scores • Baseline AN count • Geographic distribution	The subgroups are defined as follows: • Randomization stratification variables – Hurley Stage (II, III) – Concomitant antibiotic therapy (Yes, No) – Anti-TNF-α treatment (Treatment naïve, Previous treatment) • Sex (Male, Female) • BMI category: Normal (< 25 kg/m^2) , Overweight (25 – < 30 kg/m^2), obese (30 – < 40 kg/m^2), Morbid obesity (≥ 40 kg/m^2) • Age at diagnosis of HS (< median, ≥ median) • Age at study entry (<65 years vs. ≥ 65 years) • Duration of HS (< median, ≥ median) • Race (White vs. None-White)

		• Baseline AN count (< 5, 5-10, >10) • Baseline inflammatory nodule count (< median, ≥ median) • Baseline abscess count (< median, ≥ median) • Baseline draining fistula count (< median, ≥ median) • Baseline modified Sartorius score (< median, ≥ median) • Previous systemic HS treatment (includes TNF inhibitor use but excludes antibiotic use) • Geographic distribution (Northeastern Region, Southeastern Region, Western Region, Midwestern Region, and Southwestern Region)

Sponsor Approval Form

Project: BZA52227 Sponsor: ChemoCentryx, Inc.

Date: 10/07/2020 Protocol: CL016_168

Item(s) Submitted for Approval	File name (for electronic documents)
☐ * System Configuration	
☐ * Rights and Roles	
☐ * eCRF Design – PDF of forms	
☐ * Edit Check Specifications	
☐ eCRF Completion Guidelines	
☐ SAE Reconciliation Guidelines	
☐ Coding Report	
☐ Randomization Schedule Specification	
☐ Randomization Plan	
☐ Unblinding Plan	
☐ Study Data Transfer Requirements	
☐ SDTM Specifications	
☐ SDTM Transfer	
☐ CDF Specifications	
☐ CDF Test Transfer	
☐ Site Payment Criteria Specifications	
☒ Statistical Analysis Plan – text	G:\Clients\ChemoCentryx\BZA52227\Bios\Statistical Analysis Plan\Drafts\v7.0\CCX_168 SAP 07OCT2020
☐ Statistical Analysis Plan – data display shells	
☐ Statistical Reporting – TLF Delivery	
☐ Statistical Report/Summary	
☐ Other, specify:	

Signature indicates sponsor approved as delivered, unless modifications are listed/referenced below:

- [None]

*For approvals of documents with changes post-production, please select one of the following methods to verify that the changes have been completed successfully:

☐ IQVIA Biotech performs QC and documents successful implementation of changes. (Standard practice)

☐ In addition to IQVIA Biotech QC, Sponsor will verify changes before changes are implemented in production:

☐ With proof of changes (i.e. screen shots or annotated PDFs).

☐ With User Acceptance Testing (UAT) (May affect timeline and/or cost)

Signatures:

PPD

Principal Biostatistician
IQVIA Biotech

PPD

Date

PPD

PPD

Executive Director of Biostatistics
ChemoCentryx, Inc.

PPD

Date